

3/2025 dicembre

Rivista di diritto dei media

ISSN 2532-9146

DIRETTORE RESPONSABILE

Oreste Pollicino

DIRETTORI

Marco Bassini

Marina Castellaneta

Carlo Melzi d'Eril

Giulio Enea Vigevani

EDITORE

Carlo Melzi d'Eril

REDAZIONE*Cordinamento editoriale*

Rebecca Bovatti (Università di Milano)

Martina Cazzaniga (Università di Milano-Bicocca)

Giulia Napoli (Università di Milano-Bicocca)

Componenti

Flavia Bavetta (Università Bocconi – Milano)

Amalia Mariana Buzura (Avvocato in Milano)

Teresa Catalano (Università di Bari)

Federico Ceci (Università di Bari)

Marco Cecili (Università di Firenze)

Alessia Forte (Università di Milano)

Maria Galbusera (Università di Milano – Bicocca)

Giuseppe Gallo (Università di Messina)

Giovana Lopes (Tilburg University)

Matteo Monti (Università Bocconi – Milano)

Giuseppe Muto (Università Bocconi – Milano)

Federica Paolucci (Università Bocconi – Milano)

Stefania Rutigliano (Università di Bari)

SEDE

Studio legale Melzi d'Eril Vigevani

Via San Barnaba 32 - 20122 Milano

e-mail: redazione@rivistadidirittodeimedia.it

ANNO 2025 - CODICE ISSN 2532-9146

Impaginazione a cura di Chiara Pavesi

COMITATO SCIENTIFICO

Ernesto Apa <i>Avvocato in Roma</i>	Alberto Maria Gambino <i>Università Europea</i>	Giovanni Pitruzzella <i>Corte costituzionale</i>
Fabio Basile <i>Università di Milano</i>	Gianluca Gardini <i>Università di Ferrara</i>	Francesco Pizzetti <i>Università di Torino</i>
Mirzia Bianca <i>Sapienza – Università di Roma</i>	Glauco Giostra <i>Sapienza – Università di Roma</i>	Andrea Pugiotto <i>Università di Ferrara</i>
Raffaele Bifulco <i>LUISS Guido Carli</i>	Ottavio Grandinetti <i>Avvocato in Roma</i>	Margherita Ramajoli <i>Università di Milano</i>
Francesca Biondi <i>Università di Milano</i>	Enrico Grosso <i>Università di Torino</i>	Sofia Ranchordás <i>LUISS Guido Carli</i>
Carlo Blengino <i>Avvocato in Torino</i>	Antonio Gullo <i>LUISS Guido Carli</i>	Giorgio Resta <i>Università Roma Tre</i>
Manfredi Bontempelli <i>Università di Milano</i>	Antonio Iannuzzi <i>Università Roma Tre</i>	Giovanni Maria Riccio <i>Università di Salerno</i>
Elda Brogi <i>European University Institute</i>	Simone Lonati <i>Università Bocconi</i>	Gianpaolo Maria Ruotolo <i>Università di Foggia</i>
Giuseppe Busia <i>ANAC</i>	Nicola Lupo <i>LUISS Guido Carli</i>	Lucia Scaffardi <i>Università di Parma</i>
Licia Califano <i>Università di Urbino, già GPD</i>	Enrico Maria Mancuso <i>Università Cattolica</i>	Salvatore Scuto <i>Avvocato in Milano</i>
Angelo Marcello Cardani <i>Università Bocconi, già AGCOM</i>	Vittorio Manes <i>Università di Bologna</i>	Sergio Seminara <i>Università di Pavia</i>
Marta Cartabia <i>Università Bocconi</i>	Michela Manetti <i>Università di Siena</i>	Salvatore Sica <i>Università di Salerno</i>
Corrado Caruso <i>Università di Bologna</i>	Roberto Mastroianni <i>Università di Napoli “Federico II”</i>	Pietro Sirena <i>Università Bocconi</i>
Massimo Ceresa-Gastaldo <i>Università Bocconi</i>	Alessandro Melchionda <i>Università di Trento</i>	Bruno Tonoletti <i>Università di Pavia</i>
Adolfo Ceretti <i>Università di Milano – Bicocca</i>	Luigi Montuori <i>GDPA</i>	Luisa Torchia <i>Università Roma Tre</i>
Marco Cuniberti <i>Università di Milano</i>	Donatella Morana <i>Università di Roma Tor Vergata</i>	Gianluca Varraso <i>Università Cattolica</i>
Marilisa D’Amico <i>Università di Milano</i>	Antonio Nicita <i>LUMSA, già AGCOM</i>	Francesco Viganò <i>Corte costituzionale</i>
Giuseppe De Marzo <i>Suprema Corte di Cassazione</i>	Marco Orofino <i>Università di Milano</i>	Luciano Violante <i>Futuri Probabili</i>
Filippo Danovi <i>Università di Milano – Bicocca</i>	Ilia Pasquali Cerioli <i>Università di Milano</i>	Lorenza Violini <i>Università di Milano</i>
Filippo Donati <i>Università di Firenze</i>	Paolo Passaglia <i>Università di Pisa</i>	Roberto Zaccaria <i>Università di Firenze</i>
Michele Faioli <i>Università Cattolica</i>	Davide Petrini <i>Università di Torino</i>	Nicolò Zanon <i>Università di Milano</i>
Giusella Finocchiaro <i>Università di Bologna</i>	Vincenzo Pezzella <i>Suprema Corte di Cassazione</i>	Vincenzo Zeno-Zencovich <i>Università Roma Tre</i>
Maurizio Fumoglia <i>Suprema Corte di Cassazione</i>	Marina Pietrangelo <i>ITTIG-CNR</i>	

COMITATO DEGLI ESPERTI PER LA VALUTAZIONE

Maria Romana Allegri <i>Sapienza – Università di Roma</i>	Federica De Simone <i>Università della Campania</i>	Costanza Nardocci <i>Università di Milano</i>
Giulio Allevato <i>IE University</i>	Gabriele Della Morte <i>Università Cattolica</i>	Stefania Ninatti <i>Università di Milano – Bicocca</i>
Benedetta Barbisan <i>Università di Macerata</i>	Luca Di Majo <i>Università della Campania</i>	Piergiuseppe Otranto <i>Università di Bari</i>
Marco Bellezza <i>Dip. per la trasformazione digitale</i>	Federica Fabrizzi <i>Sapienza – Università di Roma</i>	Omar Makimov Pallotta <i>Università di Macerata</i>
Marco Biasi <i>Università di Milano</i>	Fernanda Faini <i>Università Pegaso</i>	Anna Papa <i>Università di Napoli “Parthenope”</i>
Daniela Bifulco <i>Corte di cassazione</i>	Fabio Ferrari <i>Università di Verona</i>	Francesca Paruzzo <i>Università di Torino</i>
Elena Bindi <i>Università di Siena</i>	Roberto Flor <i>Università di Verona</i>	Irene Pellizzone <i>Università di Milano</i>
Paolo Bonetti <i>Università di Milano – Bicocca</i>	Giulia Formici <i>Università di Parma</i>	Sabrina Peron <i>Avvocato in Milano</i>
Monica Bonini <i>Università di Milano – Bicocca</i>	Federico Furlan <i>Università di Milano – Bicocca</i>	Federico Gustavo Pizzetti <i>Università di Milano</i>
Fernando Bruno <i>AGCOM</i>	Alessandra Galluccio <i>Università di Milano</i>	Benedetto Ponti <i>Università di Perugia</i>
Daniele Butturini <i>Università di Verona</i>	Giovanni Battista Gallus <i>Avvocato in Cagliari</i>	Augusto Preta <i>International Institute of Communications</i>
Simone Calzolaio <i>Università di Macerata</i>	Emma Imparato <i>Università di Napoli “L’Orientale”</i>	Federico Riboldi <i>Avvocato in Milano</i>
Quirino Camerlengo <i>Università di Pavia</i>	Elisabetta Lamarque <i>Università di Milano – Bicocca</i>	Francesca Rosa <i>Università di Urbino</i>
Nicola Canzian <i>Università di Milano – Bicocca</i>	Erik Longo <i>Università di Firenze</i>	Marianna Russo <i>Università della Campania</i>
Marina Caporale <i>Università di Modena-Reggio Emilia</i>	Valerio Lubello <i>Università Bocconi</i>	Lucia Sciannella <i>Università di Teramo</i>
Andrea Cardone <i>Università di Firenze</i>	Claudia Massa <i>Università di Napoli “Federico II”</i>	Monica Alessia Senor <i>GDPD</i>
Bruno Carotti <i>LUISS Guido Carli</i>	Antonio Manganelli <i>Università di Siena</i>	Silvio Sonnatì <i>Università Bocconi</i>
Stefano Catalano <i>Università di Verona</i>	Paola Marsocci <i>Sapienza – Università di Roma</i>	Stefania Stefanelli <i>Università di Perugia</i>
Pier Giorgio Chiara <i>Università di Bologna</i>	Claudio Martinelli <i>Università di Milano – Bicocca</i>	Giulia Tiberi <i>Università dell’Insubria</i>
Salvatore Curreri <i>Università di Enna “Kore”</i>	Alberto Mattiacci <i>Sapienza – Università di Roma</i>	Luca Vanoni <i>Università di Milano</i>
Monica Delsignore <i>Università di Milano – Bicocca</i>	Massimiliano Mezzanotte <i>Università di Teramo</i>	Scilla Vernile <i>Università di Milano – Bicocca</i>
Giovanni De Gregorio <i>Católica Global School of Law</i>	Francesco Paolo Micozzi <i>Avvocato in Cagliari</i>	Pietro Villaschi <i>Università di Milano</i>
Giovanna De Minico <i>Università di Napoli “Federico II”</i>	Lorenzo Nannipieri <i>CNR-ITTIG</i>	Paolo Zicchittu <i>Università di Milano – Bicocca</i>

INTERNATIONAL ADVISORY BOARD

Shulamit Almog
University of Haifa

Max Baumgart
Universität Luzern

Edoardo Celeste
Dublin City University

Agustina Del Campo
*Center for Studies on Freedom of Expression
and Access to Information*

Elaine Fahey
City Law School, University of London

Irene Kamara
Tilburg University

Uta Kohl
University of Southampton

Krystyna Kowalik-Bańczyk
Tribunale dell'Unione europea

Michale Geist
University of Ottawa

Thomas Margoni
KU Leuven

Christopher Marsden
Monash University

Manuel D. Masseno
Instituto Politécnico de Beja

Dawn C. Nunziato
George Washington University

Miquel Pequera
Universitat Oberta de Catalunya

Marius Dragomir
Central European University

Andrej Savin
Copenhagen Business School

Bilyana Petkova
University of National and World Economy

Thomas Wischmeyer
Universität Bielefeld

MediaLaws - Rivista di diritto dei media è una rivista quadrimestrale telematica, ad accesso libero, che si propone di pubblicare saggi, note e commenti attinenti al diritto dell'informazione italiano, comparato ed europeo.

La rivista nasce per iniziativa di Oreste Pollicino, Giulio Enea Vigevani, Carlo Melzi d'Eril e Marco Bassini e raccoglie le riflessioni di studiosi, italiani e stranieri, di diritto dei media.

I contributi pubblicati sulla *Rivista di diritto dei media* sono redatti dagli autori e concessi a titolo gratuito, senza oneri per gli stessi.

Tutti i contributi sono pubblicati sotto licenza “Creative Commons Attribuzione – Non commerciale 4.0 International (CC BY-NC 4.0)”. La licenza consente la riproduzione, la distribuzione, la comunicazione al pubblico e la rielaborazione dei contenuti per scopi leciti e non commerciali, a condizione che sia adeguatamente attribuita la paternità dell'opera e indicata la fonte originale secondo i termini previsti dalla licenza.

Per quanto non espressamente disciplinato dalla licenza, restano applicabili le normative vigenti in materia di diritto d'autore e dei diritti connessi.

La rivista aderisce al *Code of Conduct and Best Practice Guidelines for Journal Editors* elaborato dal Committee on Publication Ethics (COPE).

La qualità e il rigore scientifico dei contributi sono garantiti mediante procedura di valutazione differenziate in relazione alla tipologia di contributi, come specificato nel Regolamento, che sono affidate a un comitato di esperti selezionati secondo criteri di competenza, rotazione e aggiornamento periodico.



MediaLaws - Rivista di diritto dei media
Regolamento per la pubblicazione dei contributi

1. “Rivista di diritto dei media” è una rivista telematica e ad accesso aperto che pubblica con cadenza quadrimestrale contributi attinenti al diritto dell’informazione e dei media, con una particolare attenzione all’impatto delle tecnologie emergenti. È facoltà della Direzione programmare la pubblicazione di numeri speciali di carattere monografico in aggiunta ai tre numeri annuali rilasciati con cadenza quadrimestrale.
2. Gli organi della rivista sono la Direzione, il Comitato scientifico, il Comitato degli esperti per la valutazione, l’International advisory board e la Redazione. L’elenco dei componenti dei predetti organi è pubblicato sul sito della rivista e su ogni fascicolo. L’elenco dei componenti del Comitato scientifico è sottoposto ad aggiornamento con cadenza quadriennale. L’elenco dei membri del Comitato degli esperti per la valutazione e dell’International advisory board è sottoposto ad aggiornamento con cadenza annuale.
3. Ciascun numero della rivista si compone delle seguenti sezioni: “Saggi”, “Note”, “Cronache e commenti” e “Recensioni”. I singoli numeri potranno altresì ospitare, in via d’eccezione, contributi afferenti a sezioni diverse (per esempio, “Editoriali”).
4. La sezione “Saggi” ospita contributi che trattano in maniera estesa e approfondita un tema di ricerca, con taglio critico e supporto bibliografico. I contributi pubblicati nella sezione “Saggi” sono normalmente di lunghezza non inferiore alle 8.000 parole e non superiore alle 20.000 parole.
5. La sezione “Note” ospita commenti alle novità giurisprudenziali provenienti dalle corti italiane, europee e straniere, oltre che a provvedimenti di autorità indipendenti con un impatto rilevante sulle tematiche della rivista. I contributi destinati alla sezione “Note” sono normalmente di lunghezza compresa tra le 4.000 e 10.000 parole.
6. La sezione “Cronache e commenti” ospita commenti a questioni e novità giuridiche di attualità nella dimensione nazionale, europea e comparata. I contributi destinati alla sezione “Cronache” sono normalmente di lunghezza compresa tra le 2.000 e 10.000 parole.
7. La sezione “Recensioni” ospita segnalazioni e commenti relativi a opere monografiche o collettanee di recente pubblicazione.
8. La richiesta di pubblicazione di un contributo è inviata all’indirizzo di posta elettronica redazione@rivistadidirittodeimedia.it, corredata dei dati, della qualifica e dei recapiti dell’autore, nonché della dichiarazione che il contributo sia esclusiva opera dell’autore e, nel caso in cui lo scritto sia già destinato a pubblicazione, l’indicazione della sede editoriale.
9. Affinché ciascun contributo possa essere valutato preliminarmente dalla Direzione, l’autore deve inviare una versione conforme alle regole editoriali della rivista pubblicate sul sito, un abstract in lingua italiana e inglese e un elenco di cinque parole chiave. Il mancato rispetto dei criteri redazionali costituisce motivo di rigetto della proposta.
10. La Direzione effettua un esame preliminare del contributo, verificando l’attinenza con i temi trattati dalla rivista e il rispetto dei requisiti minimi della pubblicazione.
11. In caso di esito positivo, la Direzione procede ad assegnare il contributo alla sezione opportuna e a sottoporlo alla procedura di valutazione scientifica (c.d. “referaggio”). In caso di esito negativo (c.d. “*desk rejection*”), la Direzione informa l’autore del contributo del rigetto.

12. I contributi destinati alla sezione “Saggi” sono inviati alla valutazione, secondo il metodo del doppio cieco, di due revisori, di cui almeno uno scelto dall’elenco degli esperti per la valutazione della rivista secondo il criterio della competenza, della conoscenza linguistica e della rotazione. I contributi destinati alle sezioni “Note”, “Cronache e commenti” e “Recensioni” sono sottoposti alla valutazione di un revisore anonimo, tendenzialmente scelto tra i membri del Comitato degli esperti per la valutazione nel rispetto dei criteri sopra ricordati. A discrezione della Direzione, i contributi in lingua straniera che presentano un taglio internazionale possono essere affidati alla valutazione di membri dell’International advisory board. I revisori ricevono una scheda di valutazione, da consegnare compilata alla Direzione entro il termine da essa indicato. Nel caso di tardiva o mancata consegna della scheda, la Direzione si riserva la facoltà di scegliere un nuovo revisore. La Direzione garantisce l’anonimato della valutazione.

13. I revisori formulano un giudizio del contributo sottoposto alla loro valutazione scegliendo tra quattro diverse opzioni: “contributo pubblicabile”, quando il lavoro possa essere pubblicato nella sua veste attuale, senza necessità di revisioni anche minori; “contributo pubblicabile con revisioni non sostanziali”, nel caso in cui il contributo sia ritenuto pubblicabile soltanto all’esito di alcune modifiche non impattanti sulla struttura e sul contenuto del lavoro e senza necessità di un’ulteriore valutazione; “contributo pubblicabile con revisioni sostanziali”, nel caso in cui il contributo sia ritenuto pubblicabile sotto condizione di modifiche di rilevante entità che rendono necessario un nuovo giudizio da parte degli stessi revisori ovvero da parte di un revisore terzo; “contributo non pubblicabile”, nel caso in cui il lavoro non presenti caratteristiche che lo rendano pubblicabile anche all’esito di eventuali modifiche sostanziali. In ogni caso, la Direzione ha facoltà di formulare un giudizio complessivo alla luce delle osservazioni preliminari da essa formulate in merito al contributo e in considerazione dell’esito delle procedure di valutazione, specialmente nel caso di giudizio divergente da parte di due revisori. È facoltà della Direzione, in caso di divergenze di giudizio, richiedere una terza valutazione ad altro esperto.

14. Una volta ricevuta comunicazione dell’esito della procedura di valutazione, in caso di giudizio positivo ovvero positivo subordinato a modifiche (sostanziali e non sostanziali), l’autore dovrà inviare una versione definitiva ed eventualmente aggiornata del contributo. Nel caso di giudizio positivo subordinato a revisioni non sostanziali, l’autore dovrà inviare il testo definitivo in modalità revisioni in modo da permettere l’individuazione delle modifiche apportate. In caso di giudizio positivo subordinato a revisioni sostanziali, l’autore dovrà inviare il testo definitivo in modalità revisioni e una scheda recante una dettagliata spiegazione delle modifiche effettuate rispetto a ciascuno dei rilievi formulati dai valutatori. Nel caso di valutazione negativa, non è consentito all’autore sottoporre nuovamente il contributo alla rivista.

15. Le valutazioni vengono archiviate dalla Redazione per almeno tre anni.

16. A discrezione della Direzione, i saggi di autori di particolare autorevolezza o richiesti dalla direzione possono essere pubblicati senza essere sottoposti alla procedura di referaggio a doppio cieco ovvero essere sottoposti a mero referaggio anonimo. Inoltre, la Direzione potrà eccezionalmente richiedere la valutazione di due revisori per contributi destinati alle sezioni “Note” e “Cronache e commenti”.

17. La Direzione ha facoltà di disporre a suo insindacabile giudizio la pubblicazione in forma di anticipazioni di contributi che abbiano positivamente terminato la procedura di valutazione e siano stati inviati in versione definitiva dall’autore in anticipo sulla data prevista di pubblicazione del rispettivo numero. In ogni caso, l’invio del contributo e la valutazione positiva alla pubblicazione non costituiscono titolo per la pubblicazione dello stesso in forma di anticipazione, che è rimessa alla decisione della Direzione in base a fattori quali l’attualità del lavoro e i tempi di pubblicazione e lavorazione del numero in formazione.

Sezione monografica **Intelligenza artificiale, decisione pubblica e stati di emergenza al tempo dello stato digitale**

- 14 Introduzione**
- 15 Intelligenza artificiale e processo decisionale. Il principio di trasparenza in alcune note di comparazione**
Emma A. Imparato
- 26 Trasparenza algoritmica e *accountability* nell'emergenza pandemica. Un'analisi comparata tra Unione europea, Stati Uniti e Cina**
Lucia G. Sciannella
- 61 Tra efficienza e tutele: le decisioni amministrative automatizzate in prospettiva comparata**
Angela Ferrari Zumbini, Martina Conticelli, Marta Infantino
- 84 AI e amministrazione federale statunitense: strategie, cambi di rotta e implicazioni globali**
Laura Fabiano
- 102 Tecnologie emergenti e Stati di emergenza: una comparazione tra Belgio e Israele**
Giacomo Giorgini Pignatiello
- 125 Algorithmic Discrimination in States of Emergency: The Role of EU Law in Promoting Counter-Strategies**
Gabriele Rugani

Saggi

- 150 The Brussels Sphinx's Riddle. What is a high-risk AI System?**
Andrea Bertolini, Federica Fedorczyk, Marta Mariolina Mollicone, Guilherme Migliora
- 197 Rule of Law in the Age of AI**
Federica Paolucci, Oreste Pollicino, Charlotte Waldron
- 247 Il lato oscuro dell'Intelligenza artificiale: come proteggere i dati personali dei lavoratori?**
Marco Capece
- 258 La convenzione ONU sul crimine informatico: tra obiettivi globali e tensioni geopolitiche**
Elisabetta Pino

Note

- 274 Transmission of Intercepted Data Obtained in a Criminal Investigation to the Competition Authority: The Grand Chamber Judgment in the Ships Waste Oil Case**
Claudia Calabrese
- 294 Diritto di critica e sanzione detentiva: "*chilling effect*" anche al di fuori della professione giornalistica?**
Alice Tadiotto
- 311 In tema di responsabilità del *blogger* per il contenuto diffamatorio. Nota a Cass. civ., sez. III, 27 giugno 2025, n. 17360**
Nicolò Zammarchi
- 323 L'indebito nella gestione collettiva dei diritti d'autore: limiti dell'autotutela economica e responsabilità dell'intermediario**
Edoardo Salvatore

Cronache

335 La stretta su intercettazioni e ordinanze

**cautelari: la stampa non più cane da
guardia della democrazia ma sciacallo
che si nutre di pettegolezzo**

Carlo Melzi d'Eril

348 La Dichiarazione per la sovranità

**digitale europea e l'utilità (trascurata)
del ricorso alla geopolitica**

Paolo Passaglia

359 L'uso della forza nel dominio cibernetico

ai sensi del diritto internazionale

Rita Caracuzzo

Focus

The Venice Commission Doctrine on Digital Technologies and the Rule of Law

14 Introduction

15 Artificial Intelligence and Decision-Making: The Principle of Transparency in a Comparative Perspective

Emma A. Imparato

26 Algorithmic Transparency and Accountability in the Pandemic Emergency: a Comparative Analysis of the European Union, the United States, and China

Lucia G. Sciannella

61 Between Efficiency and Due Process: Automated Administrative Decisions in a Comparative Perspective

Angela Ferrari Zumbini, Martina Conticelli,
Marta Infantino

84 Artificial Intelligence and the U.S. Federal Administration: Strategies, Policy Shifts, and Global Implications

Laura Fabiano

102 Emerging Technologies and States of Emergency: A Comparative Analysis of Belgium and Israel

Giacomo Giorgini Pignatiello

125 Algorithmic Discrimination in States of Emergency: The Role of EU Law in Promoting Counter-Strategies

Gabriele Rugani

Essays

150 The Brussels Sphinx's Riddle. What is a high-risk AI System?

Andrea Bertolini, Federica Fedorczyk, Marta
Mariolina Mollicone, Guilherme Migliora

197 Rule of Law in the Age of AI

Federica Paolucci, Oreste Pollicino,
Charlotte Waldron

242 The Dark Side of Artificial Intelligence: how to protect workers' personal data?

Marco Capece

258 The UN Convention on Cybercrime: Between Global Objectives and Geopolitical Tensions

Elisabetta Pino

Case notes

274 Transmission of Intercepted Data Obtained in a Criminal Investigation to the Competition Authority: The Grand Chamber Judgment in the Ships Waste Oil Case

Claudia Calabrese

294 Right to Criticise and Prison Sentence: A Chilling Effect Even for Those Who Are not Professional Journalists?

Alice Tadiotto

311 The Italian Supreme Court Rules on the Blogger's Liability for Defamatory Content

Nicolò Zammarchi

323 Undue Influence in Collective Management of Copyright: Limits of Economic Self-Defense and Intermediary Liability

Edoardo Salvatore

Comments

- 335 The Role of Journalism in the Italian Legal Order: From a Watchdog Function to Gossip-driven Sensationalism?**

Carlo Melzi d'Eril

- 348 The Declaration for European Digital Sovereignty and the Added Value of Geopolitics**

Paolo Passaglia

- 359 The Use of Force in the Cyber Domain Under International Law**

Rita Caracuzzo

Sezione monografica

**Intelligenza artificiale, decisione pubblica
e stati di emergenza al tempo
dello stato digitale**

Introduzione

La sezione monografica che qui si presenta, intitolata *Intelligenza artificiale, decisione pubblica e stati di emergenza al tempo dello stato digitale*, realizzata nell'ambito del PRIN 2020 "*Decision-making in the Age of Emergencies*" dall'unità dell'Università di Napoli "L'Orientale"¹, intende offrire una riflessione critica su uno dei nodi centrali del diritto contemporaneo: l'impatto delle tecnologie algoritmiche sull'esercizio della funzione pubblica e, in particolare, sulla gestione delle situazioni emergenziali. Negli ultimi anni, il ricorso a strumenti digitali e a sistemi di intelligenza artificiale (IA) si è intensificato in modo significativo, fino a divenire un tratto distintivo della trasformazione in atto delle amministrazioni pubbliche e delle modalità di esercizio del potere statale. L'accelerazione prodotta dalle recenti crisi globali – dalla pandemia da Covid-19 alle sfide poste dai mutamenti climatici, fino ai rischi di natura geopolitica e securitaria – ha reso evidente come le emergenze costituiscano non soltanto un banco di prova per gli ordinamenti costituzionali, ma anche un terreno privilegiato per la sperimentazione di nuove tecniche di governo fondate sulla raccolta massiva dei dati, sulla loro elaborazione algoritmica e sull'uso di strumenti di supporto decisionale.

In tale contesto, la riflessione giuridica comparata si rivela essenziale. Le esperienze maturate nei diversi ordinamenti mostrano approcci eterogenei rispetto all'integrazione degli strumenti di IA nei processi decisionali pubblici, mettendo in luce sia opportunità di razionalizzazione e di efficienza, sia rischi connessi alla tenuta dei principi fondamentali dello Stato di diritto. La comparazione consente, da un lato, di individuare tratti comuni nell'evoluzione delle pratiche amministrative digitalizzate; dall'altro, di cogliere le differenze derivanti dalle tradizioni costituzionali, dai modelli di amministrazione e dalle culture giuridiche che plasmano l'uso dell'IA nelle decisioni pubbliche. In particolare, emerge il rischio che le situazioni di emergenza, intese come momenti di eccezione rispetto all'ordinario funzionamento delle istituzioni, diventino lo spazio privilegiato per introdurre strumenti tecnologici altamente invasivi, destinati poi a consolidarsi anche al di fuori della contingenza emergenziale.

Questo spazio monografico affronta tali questioni con un taglio prevalentemente di diritto pubblico comparato, esplorando il rapporto tra tecnologie digitali, principi costituzionali e garanzie democratiche. Vengono indagati i profili problematici legati al principio di legalità e di trasparenza, al rispetto dei diritti fondamentali – in particolare alla protezione dei dati personali e al divieto di discriminazioni – e al ruolo imprescindibile della supervisione umana. Particolare attenzione è dedicata al modo in cui gli ordinamenti stanno reagendo alle sfide poste dal cosiddetto stato digitale, nel quale la legittimità della decisione pubblica non dipende più soltanto dalla conformità alle regole procedurali e sostanziali, ma anche dall'affidabilità tecnica e dall'intelligibilità degli algoritmi che la orientano.

Il contributo dei diversi autori qui raccolti intende offrire strumenti teorici e pratici per valutare l'impatto dell'IA sulle decisioni pubbliche e, più in generale, sul rapporto tra emergenza e ordinamento. L'obiettivo è duplice: da un lato, contribuire alla comprensione di un fenomeno che sta rapidamente ridisegnando le categorie tradizionali del diritto costituzionale e amministrativo; dall'altro, sollecitare una riflessione critica sul rischio di una "normalizzazione" dell'eccezione tecnologica, nella quale la logica emergenziale diventa il pretesto per introdurre forme di governo algoritmico difficilmente reversibili.

Con questo spirito, la sezione monografica si propone non solo come spazio di analisi e confronto accademico, ma anche come occasione per interrogarsi sul futuro delle nostre democrazie costituzionali di fronte all'avanzare dello stato digitale.

Emma A. Imparato
Lucia G. Sciannella
Giacomo Giorgini Pignatiello

¹ PRIN 2020 "Decision-Making in the Age of Emergencies. New Paradigms in Recognition and Protection of Rights [DeMa]" – CUP C61B22002230001.

Intelligenza artificiale e processo decisionale. Il principio di trasparenza in alcune note di comparazione*

Emma A. Imparato

Sommario

1. Profili introduttivi. 2. La normativa nell'ordinamento francese. – 2.1 Il caso portoghese. – 3. Alcune pronunce giurisprudenziali. - 3.1. L'ipotesi lituana. – 4. Brevi osservazioni finali.

1. Profili introduttivi

L'uso di strumenti di IA desta da sempre una certa apprensione. I timori ampiamente espressi dalla dottrina nonché dai giudici¹, mossi da diverse ragioni, sembrano trovare la causa più rilevante nell'ipotesi del processo decisionale e, in particolare, rispetto al livello di influenza che l'uso della intelligenza artificiale può esercitare sul processo legato alla decisione umana nonché riguardo all'eventuale conseguente 'rimpiazzo' degli attori processuali classici - cioè degli esseri umani - con entità meccaniche e virtuali². Questa ipotesi della sostituzione degli umani come decisori se sembra poi vedere talora un'ulteriore apprensione per quanto riguarda l'osservanza di tutta una serie di principi fondamentali vigenti in particolare nel diritto amministrativo, talvolta, per altro verso, porta pure taluni studiosi ad assumere una posizione ottimistica. Ritenendosi che i sistemi computerizzati possano offrire un importante contributo alla realizzazione di una decisione razionale, senza arbitrio, soprattutto nel caso della giustizia, si pensa che si possa garantire così una maggiore imparzialità.

In ogni caso appare qui evidente la connessione tra i sistemi di IA e le procedure amministrative: quando la IA viene applicata nel contesto del processo decisionale e in particolare di tipo amministrativo, questa entra

*L'articolo è stato sottoposto, in conformità al regolamento della Rivista, a referaggio "a doppio cieco".

Il presente contributo è realizzato nell'ambito del PRIN 2020 "Decision-Making in the Age of Emergencies. New Paradigms in Recognition and Protection of Rights [DeMa]" – CUP C61B22002230001

¹ V., nel caso francese, J.B. Duclercq, *Les algorithmes en procès*, in *RFDA*, 2018, spec 13 ss.

² In questi termini, v., anche, R. Malheiro de Macedo, *Inteligência Artificial. Justiça Digital e Tribunais*, in *igc.fd.uc.pt*, 2021, 2.

nel campo del diritto amministrativo divenendo così centrale la questione di quali regole e principi del diritto amministrativo siano influenzati dall'Automated decision making-ADM e dall'IA e, viceversa, quali regole e principi del diritto amministrativo sono particolarmente rilevanti per l'ADM, in generale e l'IA in particolare.

I principi del diritto amministrativo toccati si rivelano invero diversi, andando da quello della legalità e della imparzialità a quello della non discriminazione, della proporzionalità e della efficienza, solo per citarne alcuni. Nondimeno si deve dire per un altro principio associato spesso, quale suo «complementare», all'*accountability*³, quale è quello della trasparenza⁴, considerabile elemento fondamentale di questo nuovo ordine⁵. Quest'ultimo principio mostra contorni nuovi nel caso della IA.

Come è noto, la trasparenza è legata all'obbligo delle autorità pubbliche di consentire l'accesso ai documenti ufficiali in loro possesso, senza discriminazioni, cui si connette l'obbligo di pubblicazione dei documenti ufficiali, dovendo al contempo fornire informazioni sul proprio lavoro e sulle decisioni. Da qui discende il dovere dell'autorità pubblica di fornire motivazioni nella decisione all'interessato e il diritto di quest'ultimo di essere informato della decisione come anche della causa contro di lui ai fini di potersi poi eventualmente opporre e affrontare un equo processo. La trasparenza nel caso dell'IA può perciò assumere connotazioni diverse a seconda dei principi a cui si connette, variando conseguentemente dal diritto di essere informati sull'esistenza dell'ADM al diritto a requisiti di intellegibilità e, in alcuni casi, all'accesso al codice del software, previsto in alcuni ordinamenti, come quello francese, come si vedrà, espressamente dalla normativa. In queste ipotesi le amministrazioni sono investite così dell'obbligo di comunicare se la decisione è stata adottata (o supportata) da un algoritmo, rendendo noto non solo il funzionamento dello stesso ma anche il linguaggio informatico sorgente (c.d. codice sorgente) del sistema algoritmico⁶. Il principio della trasparenza si mostra allora, sotto quest'ultimo profilo, come la «chiave per la risoluzione»⁷, soprattutto in quei casi che vedono la tecnologica arrivare alla piena automazione e a decisioni mediante algoritmi⁸. In questa ipotesi specifica l'obbligo di motivazione si traduce così in una distinzione tra la trasparenza tecnica, ossia

³ Cfr. D.R. Resai - J.A. Kroll, *Trust but verify? A guide to algorithms and the law*, in *Har. J. L. & Tech.*, 31, 2017, 7 ss.

⁴ Previsto dalla legge 7 agosto 1990, n. 241, questo è teso a rendere la pubblica amministrazione più aperta, garantendo consapevolezza dei cittadini riguardo all'azione pubblica nei loro confronti. Tra i tanti, v., in particolare, M.S. Giannini, *Genesi e sostanza dei principi generali del diritto*, in Aa. Vv., *Scritti in onore di A. Predieri*, II, Milano, 1996, 901 ss.

⁵ Così, F. Pasquale, *The Black Box Society: The Secret Algorithms That Control Money And Information*, Cambridge (MA), 2015, 17.

⁶ In questi termini, S. Sassi, *Gli algoritmi nelle decisioni pubbliche tra trasparenza e responsabilità*, in *Analisi giuridica dell'economia*, 1, 2019, 109 ss.

⁷ Così, V. Brigante, *Evolving pathways of administrative decisions. Cognitive activity and data, measures and algorithms in the changing administration*, Napoli, 2019, spec. 165.

⁸ Su questo punto specifico, v. E. Carloni, *Le intelligenze artificiali nella pubblica amministrazione e la sfida della trasparenza*, in M. Lalli (a cura di), *L'amministrazione pubblica nell'era digitale*, Torino, 2022, 45 ss.

la divulgazione del codice sorgente dell'algoritmo, e la spiegazione ovvero la giustificazione della decisione, quest'ultima a sua volta soggetta a un'ulteriore distinzione tra la spiegazione del modello e la spiegazione delle singole decisioni. In quest'ottica solo un'applicazione ex ante della trasparenza all'algoritmo può garantire i diritti di contraddittorio e partecipazione e l'obbligo di motivazione, non mostrandosi affatto sufficiente la trasparenza ex post⁹.

Alla luce di questa breve premessa, questo lavoro mirerà a individuare alcuni aspetti coinvolti nell'utilizzo degli strumenti di IA per il miglioramento dell'attività amministrativa nel necessario processo decisionale, evidenziando quei profili che si rivelano inevitabilmente sacrificati nello sviluppo della funzione amministrativa. A tal fine sarà interessante innanzitutto esaminare alcuni ordinamenti europei che adottano una specifica normativa in materia di ADM, come quello francese e portoghese, ponendo in evidenza eventuali differenze di approccio normativo. Si cercherà poi di verificare quale è la posizione che la giurisprudenza va assumendo, quanto in particolare alla responsabilità legata al controllo umano, rispetto ai casi normativi analizzati e pur in assenza di una legislazione come è nell'ipotesi lituana. Negli ultimi tempi, incominciano ad essere diverse invece le sentenze che riportano l'attenzione talora sulla trasparenza tecnica talaltra su quella della giustificazione, ora su entrambe.

2. La normativa nell'ordinamento francese

In questo nuovo ordine, il principio di trasparenza va configurandosi come un obbligo riguardante non più solo i dati ma anche l'algoritmo¹⁰. Spesso l'amministrazione finisce tuttavia per lasciare il controllo agli algoritmi senza neanche stabilirlo espressamente¹¹. Diversi sono allora gli Stati membri dell'UE che hanno deciso di intervenire e adottare una specifica legislazione in materia di ADM conformandosi al Regolamento generale sulla protezione dei dati personali dell'UE (GDPR)¹², sia pure talora non

⁹ V., in merito, L. Torchia, *Lo stato digitale: una introduzione*, Torino, 2023, 128 ss.

¹⁰ Cfr. M. Finck, *Automated Decision-Making and Administrative Law*, in P. Cane - H. Hofmann - E.C. Ip - P. Lindseth (eds.), *Oxford Handbook of Comparative Administrative Law*, Oxford, 2020, 657 ss.

¹¹ Cfr. L. Cluzel-Métayer, *Calculer: l'influence des algorithmes sur l'édiction des décisions administratives. Les méthodes en droit administratif*, Paris, 2018, spec. 8 laddove si cita l'espressione di Lessing "Code is Law" (L. Lessing, *On Liberty in Cyberspace*, Harvard, 2000).

¹² Regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio del 27 aprile 2016, in vigore da maggio 2016, direttamente applicabile dal 25 maggio 2018 e relativo alla protezione delle persone fisiche con riguardo al trattamento e alla libera circolazione dei dati personali che abroga la direttiva 95/46/CE. Per una disamina generale sulla tutela della privacy e della protezione dei dati personali, v. M. Distefano (a cura di), *La protezione dei dati personali ed informatici nell'era della sorveglianza globale, Temi scelti*, Napoli, 2017. A questo si affianca Regolamento (UE) 2024/1689 del Parlamento europeo e del Consiglio del 13.6.2024 che stabilisce regole armonizzate sull'intelligenza artificiale e modifica i regolamenti (CE) n. 300/2008, (UE) n. 167/2013, (UE) n. 168/2013, (UE) 2018/858, (UE) 2018/1139 e (UE) 2019/2144 e le direttive 2014/90/UE, (UE) 2016/797 e (UE) 2020/1828 (regolamento sull'intelligenza artificiale). Per un commento, tra i primi, v.

collegandovisi direttamente¹³. In particolare giocano un ruolo importante qui le disposizioni di cui all'art. 22 concernente il "Processo decisionale automatizzato relativo alle persone fisiche, compresa la profilazione". Ai sensi di questo articolo, par. 1, del GDPR, invero, l'interessato ha il diritto di non essere sottoposto a una decisione basata esclusivamente sul trattamento automatizzato, compresa la profilazione, che «produca effetti giuridici che lo riguardano o che incida in modo analogo significativamente sulla sua persona». Prevedendosi tuttavia poi, al paragrafo successivo, diverse eccezioni a tale divieto tra cui quella che vede la decisione «autorizzata dal diritto dell'Unione o dello Stato membro cui è soggetto il titolare del trattamento, che precisi altresì misure adeguate a tutela dei diritti, delle libertà e dei legittimi interessi dell'interessato», si consente agli Stati membri di autorizzare determinate applicazioni di ADM. È in quest'ottica che così nell'ordinamento francese si detta una normativa specifica sul digitale e sugli algoritmi¹⁴.

Le disposizioni generali sulle decisioni amministrative automatizzate nel sistema francese non risultano tuttavia integrate nella legislazione sulla protezione dei dati ma in quella sulla procedura amministrativa e in particolare nel *Code des relations entre le public et l'administration* all'art. L. 311-3-1 che si iscrive nel titolo riguardante il "diritto di accesso ai documenti amministrativi". Queste norme, conseguentemente, possono trovare applicazione in un ambito più ampio, non limitato al processo decisionale amministrativo in cui è coinvolto il trattamento dei dati personali. Si stabilisce così, in generale, che una decisione individuale presa sulla base di un trattamento algoritmico «deve includere una dichiarazione esplicita che informi la persona interessata», laddove, al contempo, devono essere comunicate dall'amministrazione alla persona interessata, se questa ne faccia richiesta, anche «le norme che definiscono tale trattamento, nonché le caratteristiche principali della sua attuazione». Alla luce del sistema francese, peraltro, anche il codice sorgente è un documento amministrativo, potendo il cittadino richiederne la comunicazione.

Le condizioni per l'applicazione di queste disposizioni sono state fissate da un decreto del *Conseil d'État* del 14 marzo 2017¹⁵.

Il principio della trasparenza, in generale ritenuto un concetto non univoco e ancor più nel caso degli algoritmi¹⁶, non sembra trovare tuttavia qui

R. Petruso - G. Smorto, *Il Regolamento europeo sull'intelligenza artificiale: una prima lettura*, in *La Nuova Giurisprudenza Civile Commentata*, 4, 2024, 989 ss. nonché, più ampiamente, G. Cassano, E. M. Tripodi (a cura di), *Il regolamento europeo sull'Intelligenza Artificiale. Commento al Reg. UE n. 1689/2024*, Rimini, 2024.

¹³ Come è nel caso della normativa svedese, *Legge sulla procedura amministrativa* (2017:900), del 28 settembre 2017, applicabile, secondo l'art. 1, al «trattamento dei casi da parte delle autorità amministrative e al trattamento delle questioni amministrative da parte dei tribunali». V., in merito, J. Wolswinkel, *Comparative study on administrative law and the use of artificial intelligence and other algorithmic systems in administrative decision-making in the member states of the Council of Europe*, Report Council of Europe, December 2022, 17.

¹⁴ Loi n° 2016-1321 del 7 ottobre 2016 per *Une République numérique*

¹⁵ Décret n° 2017-330 *relatif aux droits des personnes faisant l'objet de décisions individuelles prises sur le fondement d'un traitement algorithmique*. Sul punto v. H. Pauliat, *La décision administrative et les algorithmes: une loyauté à consacrer*, in *Revue du droit public*, 2018, 641 ss.

¹⁶ In particolare, v. Y. Paquier, *Le principe de transparence des traitements algorithmiques: de*

sufficiente garanzia¹⁷. Seppure l'amministrazione comunica alla persona interessata da una decisione individuale adottata sulla base di un trattamento algoritmico che lo richieda, in forma intelligibile e a condizione che non violi segreti protetti dalla legge, una serie di elementi informativi (tra cui, il grado e le modalità di contributo del trattamento algoritmico al processo decisionale, i dati trattati e le loro fonti, i parametri di trattamento e, se del caso, la loro ponderazione, applicati alla situazione dell'interessato nonché le operazioni eseguite dal trattamento), gli obblighi sembrano ancora troppo limitati. Si finisce così con l'assicurare il contrario dell'equità e trasparenza¹⁸. Lo stesso discorso vale per il principio di *loyauté*¹⁹ riguardante i «*concepteurs et architectes de l'intelligence artificielle*», utilizzabile affinché il rinvio all'interesse generale aiuti a superare le difficoltà di comprensione²⁰. L'obbligo di lealtà, ritenuto essenziale, non è ancora considerato tuttavia pienamente soddisfatto presupponendo l'osservanza di una serie di altri principi non previsti: la restituzione in un linguaggio chiaro che può essere compreso dagli esseri umani, «inclusi i progettisti di questi robot e quindi a maggior ragione dai cittadini e dalle autorità pubbliche, è diventata in molti casi impossibile»²¹ come tale risulta, peraltro, la previsione tutti i possibili *bias* che possono influenzare un campione²².

2.1. Il caso portoghese

Un altro interessante esempio di Paese che tenta di normare il trattamento automatizzato è anche quello del Portogallo che, in tempi piuttosto recenti, ha peraltro adottato la Carta dei diritti umani nell'era digitale²³. Non limitato, come nel caso francese, a disciplinare l'ADM, questo testo dal titolo alquanto “ambizioso”²⁴ contiene diverse disposizioni giuridiche sull'amministrazione digitale riguardante anche aspetti innovativi quale è

L'étude juridique d'un enjeu démocratique, Normandie, 2021, 22.

¹⁷ Così, H. Pauliat, *La décision administrative et les algorithmes*, cit., 649.

¹⁸ In tal senso, S. Abiteboul - G. Dowek, *Le temps des algorithmes*, Paris, 2017, 146, giacché gli algoritmi «*sont tels que nous les réalisons*».

¹⁹ Su questo principio, anche per il caso italiano, v. F. Laviola, *Algoritmico, troppo algoritmico: decisioni amministrative automatizzate, protezione dei dati personali e tutela delle libertà dei cittadini alla luce della più recente giurisprudenza amministrativa*, in *Biolaw Journal*, 3, 2020, 32 ss.

²⁰ Così J.M. Sauvé, *Le juge administratif et l'intelligence artificielle*, Discours prononcé lors de la conférence des présidents des juridictions administratives, reperibile all'url conseil-etat.fr, laddove si osserva che «*Ce principe permettrait par ailleurs de surmonter les difficultés d'intelligibilité que j'évoquais précédemment en imposant aux opérateurs numériques de se comporter conformément à l'intérêt général*».

²¹ Rapport public du Conseil d'Etat 2017, p. 110.

²² Così, O. Sibony, *Big data: des données pas si transparentes*, in *Contrepoints.org*, 2017.

²³ Lei n. 27/2021, del 17 maggio 2021, che approva la *Carta Portuguesa de Direitos Humanos na Era Digital*. Per un'attenta ricostruzione riguardante l'evoluzione storica della Carta e i suoi contenuti, v. D. Soares Farinho, *La Carta Portuguesa de Derechos Humanos en la Era Digital. Una valoración jurídica*, in *Revista Española de la Transparencia*, 2021, 85 ss.

²⁴ I questi termini J. Pereira da Silva, *Direitos Fundamentais da Era Digital*, Fundação Francisco Manuel Dos Santos, Lisboa, 2024, 50 ss.

quello relativo allo *Uso da inteligência artificial e de robôs* di cui all'art. 9.

In quest'ottica, se si può dire che in generale manca una normativa dedicata alle applicazioni di IA, nello specifico caso lusitano, pur senza offrirne alcuna definizione²⁵, ben si ha riguardo all'IA prescrivendosi che il suo uso deve essere guidato dalla tutela dei diritti fondamentali. Si deve anche garantire il «giusto equilibrio tra i principi di chiarezza, sicurezza, trasparenza e responsabilità, tenendo conto delle circostanze di ogni singolo caso e stabilendo procedure volte a evitare pregiudizi e forme di discriminazione». Nel rispetto della Costituzione portoghese e specificamente dell'art. 35 sull'uso delle tecnologie informatiche che introduce il diritto di accesso del cittadino a tutti i dati informatici che lo riguardano, si prevedono inoltre diversi compiti pubblici relativi al cyberspazio quali già stabiliti nella stessa carta costituzionale «duplicandoli», ad avviso di parte della dottrina, inutilmente²⁶.

In ogni caso, al fine di garantire un ambiente digitale che promuova e difenda i diritti umani, si stabilisce in maniera chiara la responsabilità dello Stato di promuovere «l'uso autonomo e responsabile di Internet e il libero accesso alle tecnologie dell'informazione e della comunicazione» mentre la trasparenza, oltre all'uguaglianza, è garantita espressamente soltanto per il dominio internet portoghese PT (art. 3).

3. Alcune pronunce giurisprudenziali

A fronte di una normativa specifica come anche in sua assenza, il giudice chiamato a risolvere casi concernenti procedimenti amministrativi robotizzati con una conseguente adozione di decisione algoritmica, può assumere come riferimento, oltre alla disciplina interna, quella contenuta nel Gdpr. A questa si accorda generalmente il valore ed efficacia propria dei principi generali²⁷ arrivandosi così a riconoscere innanzitutto il diritto di decifrare la logica dell'algoritmo²⁸.

Sull'accesso all'elaborazione algoritmica e al codice sorgente si rivela peraltro importante un caso francese esaminato dal *Conseil d'Etat* nel 2019²⁹, preceduto da un'altra pronuncia del 2018, questa volta adottata dal *Conseil constitutionnel* a seguito di una *saisine* relativa alle modifiche apportate alla *Loi informatique et libertés* al fine di adeguare la legge francese in materia di protezione dati al regolamento (UE)³⁰. In quest'ultima ipotesi diverse

²⁵ Così, J. Wolswinkel, *Comparative study on administrative law*, cit., 18.

²⁶ Cfr., D. Soares Farinho, *La Carta Portuguesa de Derechos Humanos en la Era Digital*, cit., 93.

²⁷ In tal senso, E. Carloni, *La trasparenza amministrativa e gli algoritmi*, in E. Belisario - G. Cassano (a cura di), *Intelligenza artificiale per la pubblica amministrazione. Principi e regole del procedimento amministrativo algoritmico*, Pisa, 2023, 221 ss.

²⁸ Così, D. Cardon, *Le pouvoir des algorithmes*, in *Pouvoirs*, n. 164, 2018, 65 ss.

²⁹ Conseil d'État n° 427916, 12 giugno 2019. V., in merito, C. Aguilon, *Le contrôle par le juge administratif du caractère non discriminatoire des décisions individuelles algorithmiques*, in *Annuaire international de justice constitutionnelle*, 2019, 605.

³⁰ Conseil constitutionnel, n. 2018-765 DC, 12 giugno 2018, *Loi relative à la protection des données personnelles*. In questa pronuncia, il giudice costituzionale francese ritenne che gli algoritmi c.d. «*auto-apprenant*», cioè in grado di modificare autonomamente le regole che

furono peraltro le considerazioni svolte tra cui quella che vede necessario, secondo lo stesso giudice costituzionale francese, rendere accessibili e comprensibili le decisioni amministrative algoritmiche. Nella sua decisione del 12 giugno, il Consiglio di Stato si è trovato ad affrontare invece il problema del diritto alla comunicazione dell'algoritmo di elaborazione, specificamente, delle domande esaminate nell'ambito di una procedura nazionale di pre-immatricolazione alle Università tramite una piattaforma – la *Parcoursup* – nonché il problema della comunicazione dei suoi codici sorgente non resi accessibili ai sindacati degli studenti. È sulla base della specifica normativa nazionale a cui si è fatto cenno e, in particolare, del Codice dei rapporti tra il pubblico e l'amministrazione riguardante, tra l'altro, la procedura e la trasparenza amministrativa applicabile anche all'elaborazione algoritmica che il *Conseil d'Etat* ha ribadito come, in linea generale, i principi della trasparenza debbano essere osservati anche nel caso di ADM. Ovvero, ad avviso del giudice francese devono essere rese accessibili le «informazioni relative ai criteri e alle procedure di esame delle loro domande, incluse quelle relative al trattamento algoritmo utilizzato, nonché le ragioni pedagogiche che giustificano la decisione presa», sia pure, ben si precisa, dovendole dare «solo ai candidati», ed escludendosi perciò, con rigetto del ricorso, la *Union nationale des étudiants de France* (UNEF).

Diverso è il discorso per il Portogallo.

Quest'ultimo ordinamento non vede invero tuttora l'adozione di pronunce giurisprudenziali in materia. L'utilizzo di algoritmi nel processo decisionale da parte dell'Amministrazione, quale esaminato alla luce del principio della trasparenza e del diritto di essere informati, non si rivela

applicano, non possono essere in ogni caso utilizzati come base esclusiva per una decisione amministrativa individuale, se non sono assoggettati a controllo né a convalida. Il *Conseil* giunge così, in ogni modo, fatta salva una parziale non conformità, la costituzionalità della legge relativa alla protezione dei dati personali, ritenendo che l'autorizzazione di una decisione amministrativa basata su un unico fondamento algoritmico non modifichi la sostanza dei poteri dell'amministrazione. In proposito v. J. B. Duclercq, *L'automatisation algorithmique des décisions administratives individuelles*, in *Revue du droit public*, 2019, 325, ove si critica l'approccio del giudice costituzionale francese allorché si è astenuto dall'esercitare il proprio potere di sollevare d'ufficio una questione per esaminarne la costituzionalità, «mentre è proprio l'esistenza del doppio obbligo di informazione e di divieto di ricorso all'intelligenza artificiale che lo ha convinto a non optare per una dichiarazione di censura dell'articolo». Più specificamente sul principio di trasparenza, quanto a questa pronuncia come anche su quella successiva del 2020, si rinvia a T. Douville, *Parcoursup: transparence des algorithmes locaux limitée à raison pour le Conseil constitutionnel*, in *IP/IT*, 2020, 09, 516. Nel pronuncia del 2020, la seconda in materia, il Consiglio costituzionale, chiamato ad esprimersi dal Consiglio di Stato (decisione n. 433296 e 433297 del 15 gennaio 2020) per una questione prioritaria di costituzionalità sulla conformità alla costituzione di una disposizione del Codice dell'Istruzione (art. L. 612-3, I, par. 5) che limita il diritto di accesso agli algoritmi locali implementati dalle università nell'ambito della procedura nazionale di pre-iscrizione al diploma di laurea, ha ritenuto la conformità costituzionale della normativa oggetto di ricorso. A tal fine ha formulato una riserva interpretativa per cui «le disposizioni impugnate non possono, senza violare il diritto di accesso ai documenti amministrativi, essere interpretate nel senso che esentano ciascuna istituzione dalla pubblicazione, al termine della procedura nazionale di preiscrizione e nel rispetto della vita privata dei candidati, se del caso, sotto forma di relazione, i criteri in base ai quali le domande sono state esaminate e specificando, se del caso, in che misura l'elaborazione algoritmica è stata utilizzata per effettuare tale esame» (§ 17). Décision n° 2020-834 QPC du 3 avril 2020.

ancora un caso affrontato dal *Tribunal constitucional*³¹. Interesse particolare assumono invece qui le questioni riguardanti la gestione pubblica illecita dell'IA, costituente un reato finanziario, e l'eventuale soggetto da ritenere responsabile³² nonché l'utilizzo di sistemi informatici adottati in particolare dall'autorità fiscale³³. In queste ultime ipotesi diversi sono i punti eccepiti andandosi dal rispetto dei limiti legali del pignoramento dei beni o del principio di proporzionalità nei pignoramenti automatici alla difficoltà dell'autorità fiscale di correggere gli errori, quando si inseriscono i dati nel sistema informatico. In una sentenza del 2020, il tribunale amministrativo ritiene la motivazione 'generica', essendosi utilizzati modelli di giustificazione standardizzati che non si adattano alle specificità di ciascun caso³⁴.

3.1. L'ipotesi lituana

Un caso particolarmente interessante è quello invece lituano che malgrado l'assenza di una legislazione generale, come è presente non solo nell'ordinamento lusitano ma anche in quello francese, ha visto la Corte Amministrativa Suprema affrontare diverse volte, a differenza del Portogallo, le problematiche connesse all'uso delle decisioni automatizzate. Il richiamo è spesso al diritto internazionale ed europeo sul trattamento dei dati personali, osservandosi che anche alla luce di questa normativa come anche della giurisprudenza UE, emerge una «maggiore necessità di tale protezione quando i dati sono trattati automaticamente»³⁵.

In quest'ottica, in una sentenza del febbraio 2020³⁶, a fronte della decisione del comune, priva di motivazione, la Corte lituana sottolinea l'irrilevanza dei mezzi ai fini della garanzia della trasparenza amministrativa. Qui il caso riguardava, in particolare, il diniego, espresso dall'*Alytus City Municipality*, di riconoscere un assegno sociale a una cittadina. L'ente comunale aveva adottato una decisione negativa in base al mero calcolo generato in modo automatizzato sulla base di dati rispetto cui gli addetti comunali - si asseriva - non disponevano di alcuna capacità tecnica per cancellarli o modificarli, essendo peraltro privi di qualunque potere di controllo della correttezza dei dati forniti dalle istituzioni. Più precisamente si sostiene che “i

³¹ V., in proposito, B. Garcia, *Limites Constitucionais da Automatização Administrativa: Considerações à Luz da Jurisprudência do Conselho de Estado*, in *Revista Portuguesa de Direito Constitucional*, 2024, 129 laddove si prendono in considerazione i due casi affrontati dal Consiglio di stato italiano nel 2019 (Cons. Stato, sez. VI, 8 aprile 2019, n. 2270 e 13 dicembre 2019, n. 8472) ai fini della mobilità degli insegnanti della scuola secondaria, per esaminarli alla luce dell'ordinamento lusitano.

³² Così J. F. F. Tavares - P. N., Costa, *Inteligência Artificial, gestão pública e controlo*, in R. Pedro - P. Caliendo (a cura di), *Inteligência Artificial no contexto do Direito Público: Portugal e Brasil*, Almedina, 2023, 73, laddove si osserva che «A “governança algorítmica” ou a “parceria humano-máquina” coloca desde logo a questão fundamental sobre quem poderá ser responsabilizado»

³³ V., in merito, L.M. Pica, *A Inteligência Artificial no Direito Tributário: fundamentos e limites constitucionais*, Almedina, 2023.

³⁴ Acórdão do Tribunal Central Administrativo Sul, 9, 2020, 6 dicembre 2020.

³⁵ Corte suprema amministrativa della lituana, A-2 280-629/2021, 9 giugno 2021.

³⁶ Corte suprema amministrativa della lituana, A-275-822/2020, 19 febbraio 2020.

mezzi scelti” dal Comune per la predisposizione di tali decisioni, nel caso specifico un sistema automatizzato, non fanno venir meno il suo «dovere di rispettare le prescrizioni delle norme giuridiche e dei principi giuridici, compresi i principi di legalità e buona amministrazione», avendo i destinatari «un’opportunità limitata di adeguare tali decisioni nella controversia che si è venuta a creare». Le ragioni addotte, tra cui il calcolo accertato in modo automatizzato, «non hanno alcun significato giuridico nella controversia», fermo restando poi in ogni caso l’obbligo della motivazione al fine di non impedire al ricorrente di comprendere le ragioni della decisione negativa adottata nei suoi confronti e quindi non svilire il suo diritto di opporsi in modo efficace e verificare se i suoi diritti sono stati lesi.

4. Brevi osservazioni finali

Alla luce della normativa ma soprattutto della giurisprudenza la trasparenza legata al potere amministrativo conferito all’autorità pubblica, come anche la delega di fatto di tali poteri alla intelligenza artificiale, si conferma un punto essenziale molto dibattuto in dottrina. Questo aspetto incontra nella visione delle corti e, in particolare, in quella lituana una posizione chiara laddove si sostiene con fermezza che l’uso della tecnologia non può esimere i funzionari pubblici dalla supervisione. Tutt’altro. Tra le righe si può leggere che si può ritenere lecito quest’uso solo qualora i funzionari possano dimostrare di aver esercitato un controllo significativo della decisione piuttosto che un semplice gesto simbolico, ben dovendo invece restare l’autorità e la competenza della pubblica amministrazione al fine di modificare la decisione, alla luce di tutti i dati rilevanti.

Nella sua giurisprudenza, la Corte Suprema Amministrativa della Lituania ha invero sottolineato, in particolare nel caso delle decisioni automatizzate, l’esigenza di garantire alcuni principi e diritti fondamentali anche nelle procedure amministrative quale il diritto a essere debitamente informati dei propri diritti e conoscere le azioni concrete nonché il diritto al contraddittorio. Orbene, alla luce di questa giurisprudenza, oltre all’obbligo di motivazione, risulta evidente un altro obbligo, ovvero quello del controllo. Gli enti pubblici, secondo i giudici, rimangono invero responsabili della “qualità” del processo decisionale basato su sistemi di automatizzazione, non essendo esentati dalla supervisione e dovendosi così superare quel “pregiudizio dell’automazione” per cui gli individui, ritenendo o presumendo i computer, in generale, e i sistemi ADM, in particolare, intrinsecamente razionali, tendono a fidarsi più delle decisioni prese dalle macchine che dagli umani, esercitando difficilmente un controllo significativo sulle decisioni assunte dall’intelligenza artificiale. Al contrario, sembra emergere chiaramente qui la necessità che resti un adeguato livello di intervento umano.

Del resto gli stessi sistemi ADM sono progettati dall’uomo e utilizzati per scopi determinati dallo stesso uomo e possono così rivelarsi difetti non identificati nel modello statistico del sistema che in circostanze specifiche può portare a decisioni errate e irrazionali. L’esempio classico citato a tal proposito è quello del sistema COMPAS della giustizia predittiva che di-

mostra, come senza volontà diretta, si possano verificare sistemi che giungono a conclusioni non previste dai loro stessi progettisti con pregiudizi discriminatori derivanti dalla costruzione degli algoritmi, dai dati utilizzati (qualità) e dalla preparazione dei dati.

Se allora anche questa, come ogni altra tecnologia, si configura alla fine come uno strumento nelle mani di chi la usa, potendosi l'IA rivelare, allorché ben costruita e utilizzata, uno strumento potente e prezioso per la società, ciò non fa venir meno l'importanza della supervisione dell'essere umano, chiamato a perfezionarla, aggiornarla e calibrarla nonché, ove necessario, correggerla.

Abstract

Questo articolo esplora l'interazione tra intelligenza artificiale (IA) e processo decisionale amministrativo, con particolare attenzione al principio di trasparenza. Lo studio evidenzia come gli strumenti di IA, quando integrati nella pubblica amministrazione, possano mettere in discussione principi fondamentali quali legalità, imparzialità, non discriminazione, proporzionalità e responsabilità. La trasparenza emerge come una questione chiave che comprende non solo l'accesso alle informazioni e al ragionamento, ma anche, in alcuni casi giurisprudenziali, la divulgazione del codice algoritmico stesso. Le prospettive comparative riguardano la Francia, il Portogallo e la Lituania, dove i diversi approcci normativi e giudiziari dimostrano sia le opportunità che le mancanze nel garantire un processo decisionale automatizzato equo e intellegibile. Sebbene l'IA possa contribuire all'efficienza e alla razionalità, l'analisi sottolinea la necessità di un controllo umano significativo per salvaguardare i diritti, prevenire i pregiudizi e garantire il giusto processo. In definitiva, la trasparenza nelle decisioni assistite dall'IA non è solo un requisito tecnico, ma una condizione fondamentale per la legittimità della governance democratica.

This paper explores the interaction between artificial intelligence (AI) and administrative decision-making, with particular attention to the principle of transparency. The study highlights how AI tools, when integrated into public administration, may challenge fundamental principles such as legality, impartiality, non-discrimination, proportionality, and accountability. Transparency emerges as a key issue, encompassing not only access to information and reasoning but also, in some jurisdictions, disclosure of the algorithmic code itself. Comparative perspectives are drawn from France, Portugal, and Lithuania, where different regulatory and judicial approaches demonstrate both opportunities and shortcomings in ensuring fair and explainable automated decision-making. While AI may contribute to efficiency and rationality, the analysis underlines the necessity of meaningful human oversight to safeguard rights, prevent bias, and uphold due process. Ultimately, transparency in AI-assisted decisions is not merely a technical requirement but a fundamental condition for legitimacy in democratic governance.

Keywords

intelligenza artificiale - trasparenza amministrativa - processo decisionale automatizzato - supervisione umana

Trasparenza algoritmica e *accountability* nell'emergenza pandemica. Un'analisi comparata tra Unione europea, Stati Uniti e Cina*

Lucia G. Sciannella

Sommario

1. Premessa. – 2. Trasparenza algoritmica e *accountability* alla prova dell'esperienza pandemica: itinerari di diritto comparato. – 3. Le certificazioni vaccinali nell'Unione europea, negli Stati Uniti e in Cina. – 3.1. Il modello europeo tra interoperabilità tecnica e deficit di *accountability* preventiva. – 3.2. Il modello statunitense tra frammentazione federale e *governance* privatistica. – 3.3. Il modello cinese: dinamismo algoritmico e opacità sistemica. – 4. Il monitoraggio algoritmico della mobilità: tre paradigmi di *governance* dei dati in emergenza. – 4.1. Dati di mobilità e prossimità “*privacy-preserving*” nel modello europeo. – 4.1.1. *Contact Tracing* e sorveglianza della mobilità: le esperienze nazionali. – 4.1.2. Dati aggregati di mobilità e trasparenza algoritmica: alcuni casi critici nei Paesi dell'UE. – 4.2. Il modello statunitense tra frammentazione federale e *surveillance laundering*. – 4.2.1. L'esperienza di *contact tracing* negli Stati Uniti tra *standard* privati e frammentazione dell'*accountability*. – 4.3. Il modello cinese tra sorveglianza sanitaria e controllo sociale. – 5. Considerazioni conclusive: verso un “costituzionalismo algoritmico dell'emergenza”?

1. Premessa

L'emergenza sanitaria innescata dalla pandemia da Covid-19 ha rappresentato un autentico “*stress test*”¹ per gli ordinamenti nazionali e per le rispettive architetture di garanzia, non soltanto per l'intensità delle limitazioni

*L'articolo è stato sottoposto, in conformità al regolamento della Rivista, a referaggio “a doppio cieco”.

Il presente contributo è realizzato nell'ambito del PRIN 2020 “Decision-Making in the Age of Emergencies. New Paradigms in Recognition and Protection of Rights [DeMa]” – CUP C61B22002230001

¹ Cfr. A. Vidaschi-L. Cuocolo, *Il Covid-19, l'ultimo stress test per gli ordinamenti democratici: uno sguardo comparato*, in *DPCE online*, 2, 2020, 1453 ss.

incidenti su diritti e libertà², ma anche per l'accelerazione – spesso senza adeguati tempi di scrutinio pubblico – di soluzioni di governo fondate su infrastrutture digitali, processi automatizzati e decisioni algoritmicamente assistite³. In effetti, in un brevissimo arco temporale, sono state adottate soluzioni tecnologiche di inedita portata – dal *contact tracing* ai modelli previsionali per le restrizioni alla mobilità, dalle certificazioni digitali come condizioni d'accesso a spazi e servizi, fino all'uso di dati di localizzazione per l'analisi dei flussi – hanno assunto una funzione materialmente allocativa dei diritti, rendendo palpabile la frizione tra efficienza dell'azione amministrativa e controllo democratico.

Ed è proprio in tale punto di tensione – tra “governo per dati” e garanzie costituzionali – che si colloca l'oggetto di indagine del presente contributo, il quale intende verificare, in chiave comparata, la tenuta dei principi di trasparenza e *accountability* nelle traiettorie di digitalizzazione emergenziale sviluppatesi nell'Unione europea, negli Stati Uniti e nella Repubblica Popolare cinese, con specifico riferimento a due snodi ad alta tensione da punto di vista costituzionale: (i) le certificazioni digitali sanitarie e (ii) i sistemi di *contact tracing*.

L'analisi comparativa non intende perseguire una mera finalità descrittiva, bensì offrire l'occasione di osservare come modelli diversi di rapporto tra potere pubblico, infrastrutture tecnologiche (spesso dipendenti da attori privati) e diritti fondamentali producano differenti criteri giustificativi, forme di controllo e disponibilità di rimedi, soprattutto allorquando l'emergenza comprime tempi e spazi di deliberazione.

L'obiettivo dell'indagine è duplice. Sul piano ricostruttivo, si intende mappare le modalità di implementazione, scelte tecnologiche e architetture di governance adottate nei tre ordinamenti durante la fase pandemica (inclusi processi di *law-making*, *procurement* e distribuzione delle responsabilità). In secondo luogo, sul piano analitico, si intende verificare se – e in che misura – tali scelte abbiano garantito l'effettività dei principi di trasparenza algoritmica e *accountability*, come autentiche categorie operative. Sotto tale specifico aspetto, per quanto concerne la trasparenza, si analizzerà la sua portata nella fase *ex ante* (valutazioni d'impatto, basi giuridiche, scelte progettuali, standard e interoperabilità) e nella fase *ex post* (accessibilità delle informazioni, tracciabilità, audit). Per ciò che concerne l'*accountability*, si valuterà la sua portata in termini di spiegabilità e controllo, di rimedi e contestabilità effettiva, nonché il grado di dipendenza infrastrutturale e contrattuale da attori privati. Una particolare attenzione sarà, da ultimo, dedicata al rischio di “stabilizzazione dell'eccezione”, ossia la trasformazione di soluzioni emergenziali in dispositivi permanenti, con ricadute durvoli sul circuito democratico.

² G. Smorto, *Soluzioni tecnologiche e lotta al Covid-19. Una riflessione critica*, in DPCE online, 1, 2023, 731 ss.; C. Perlinger, *Coronavirus e tracciamento tecnologico: alcune riflessioni sull'applicazione e sui relativi sistemi di interoperabilità dei dispositivi*, in *Actualidad Jurídica Iberoamericana*, 12bis, 2020, 836 ss.

³ Sul consolidamento di *policy* e infrastrutture *e-Health* e sul ruolo acceleratore della pandemia, cfr. G. Lofaro, *Dati sanitari e e-Health europea: tra trattamento dei dati personali e decisione amministrativa algoritmica*, in questa Rivista, 2, 2022, 200 ss.

2. Trasparenza algoritmica e *accountability* alla prova dell'esperienza pandemica: itinerari di diritto comparato

L'analisi comparata tra Unione europea, Stati Uniti e Cina offre una specifica utilità per lo studio della governance algoritmica in pandemia, in quanto consente di osservare come ordinamenti caratterizzati da diverse tradizioni costituzionali, assetti istituzionali e concezioni del rapporto tra Stato e cittadino abbiano affrontato una medesima sfida funzionale⁴, ovvero l'impiego di infrastrutture digitali e, in taluni casi, di sistemi decisionali automatizzati per perseguire finalità di salute pubblica⁵.

Al fine di evitare equivoci concettuali, l'espressione "governance algoritmica" è qui intesa in senso ampio, includendo sia decisioni algoritmicamente assistite, sia architetture tecniche (standard, protocolli, interoperabilità) che, pur non essendo "algoritmi" in senso stretto, producono effetti su accesso, circolazione e fruizione di servizi.

Sotto il profilo formale, i tre contesti esaminati presentano approcci significativamente diversi per quanto concerne il riconoscimento dei principi di trasparenza e *accountability*⁶. Nell'Unione europea, la risposta pandemica si è innestata in un quadro normativo orientato alla legalità procedurale e alla protezione dei diritti, con una tendenza a tradurre *ex ante* in vincoli giuridici e procedurali l'innovazione tecnologica. Negli Stati Uniti, l'assenza di una disciplina federale organica in materia di dati personali e di strumenti digitali per la sanità pubblica ha prodotto un panorama frammentato e policentrico, affidato a iniziative statali e settoriali, con ampi spazi di di-

⁴ Secondo un'attenta riflessione dottrinale, «la reazione dei diversi ordinamenti all'emergenza Covid-19 ha evidenziato significative e, in qualche caso, convergenti dinamiche tanto nei rapporti tra i poteri politicamente sensibili (legislativo ed esecutivo), con rilevanti conseguenze sulla forma di governo, quanto tra i livelli decisionali delle articolazioni territoriali (Federazione/Stato, regioni, municipalità), comportanti non trascurabili strascichi sull'asse verticale della forma di Stato», anche se il vero «denominatore comune delle misure decise dai governi delle democrazie avanzate, al fine di contenere la diffusione del contagio (...) è rappresentato dalla limitazione, se non addirittura, in qualche caso, dalla sospensione, di libertà e diritti costituzionalmente garantiti» (così A. Vedaschi-L. Cuocolo, *Il Covid-19, l'ultimo stress test per gli ordinamenti democratici: uno sguardo comparato*, cit., 1455).

⁵ Vi è da precisare che l'utilizzo della tecnologia in ambito sanitario rappresenta ormai una tendenza di lungo corso, soprattutto nei sistemi giuridici occidentali. Nel contesto europeo, ad esempio, l'infrastruttura di servizi digitali per la sanità elettronica, le piattaforme per la gestione di profilassi e campagne vaccinali fanno parte di una *policy* ben affermata e portata avanti da diversi anni. In tal caso, la pandemia ha funto da acceleratore, favorendo un impiego ampio e diffuso di soluzioni tecnologicamente innovative in ambito medico. Cfr., in tale ambito, P. Lombardi, *Sicurezza dei dati in ambito sanitario ed evoluzione tecnologica tra passato, presente e futuro*, in *Il diritto dell'economia*, 3, 2021, 60 ss.; G. Lofaro, *Dati sanitari e e-Health europea: tra trattamento dei dati personali e decisione amministrativa algoritmica*, cit., 200.

⁶ Sul ruolo della pandemia come "stress test" costituzionale e rivelatore di differenze strutturali, v. R. Tarchi, *L'emergenza sanitaria da COVID-19: una prospettiva di diritto comparato. Riflessioni a margine di un seminario pisano*, in Id., *L'emergenza sanitaria da COVID-19: una prospettiva di diritto comparato*, in *Rivista «Gruppo di Pisa. Dibattito aperto sul Diritto e la Giustizia Costituzionale»*, 2, 2020, 2 ss.; V. Piergigli, *Diritto costituzionale dell'emergenza*, Torino, 2023, 213 ss.

screzionalità agli operatori privati. Infine, la Repubblica Popolare cinese, dal suo canto, ha introdotto precocemente normative specifiche sugli algoritmi e sulla governance dei dati⁷, talora richiamando i medesimi principi presenti nelle normative europee, ma collocandoli entro una razionalità politico-istituzionale diversa, in cui l'asimmetria tra obblighi imposti alle piattaforme commerciali e opacità dei sistemi impiegati direttamente dal potere pubblico incide sulla funzione stessa della trasparenza e sulla disponibilità di rimedi effettivi per i destinatari delle decisioni⁸.

Sotto il profilo dell'analisi comparata, l'allargamento dello spettro d'analisi alla Repubblica Popolare Cinese risponde, inoltre, a un'ulteriore esigenza metodologica, consistente nel seguire una traiettoria di analisi che consenta di ridefinire le categorie giuridiche tradizionali alla luce di formanti tecnologici che operano come fattori ordinanti del potere. Difatti, l'adozione accelerata di dispositivi digitali per finalità di salute pubblica ha portato in superficie nuove fonti autonome di produzione del diritto – individuabili nell'insieme di scelte tecniche incorporate in codice, dataset, metriche, interfacce – che non si limitano a “eseguire” regole preesistenti, ma possono produrre effetti “quasi-normativi”, fissando soglie e classificazioni che incidono sull'accesso a diritti, servizi e libertà. Ne consegue che l'unità di analisi del comparatista non può più esaurirsi nella regola formale o nella decisione giudiziaria, ma deve includere il ciclo di vita del sistema (progettazione, regole di addestramento, implementazione, *log* e auditabilità, contestazioni, rimedi), poiché è in queste fasi che si decide la concreta traducibilità di trasparenza e *accountability* in garanzie effettive. In questa prospettiva, il formante tecnologico diventa oggetto autonomo di comparazione, utile a individuare i tratti di quella “forma di governo digitale” che l'emergenza sembra aver delineato e, in parte, normalizzato⁹.

Al fine di rendere più concreta l'analisi, l'indagine si concentrerà su due ambiti funzionali dell'ecosistema algoritmico utilizzato nella risposta pandemica: (i) le certificazioni vaccinali e, più in generale, i dispositivi di accesso condizionato a spazi e servizi; (ii) il monitoraggio della mobilità mediante dati di localizzazione e informazioni sui flussi. La scelta di tali ambiti è giustificata sia dalla pervasività che tali strumenti hanno avuto nei tre contesti, sia dall'impatto che essi hanno prodotto sulle condizioni di trasparenza, controllabilità e contestabilità delle decisioni. E, come più avanti si vedrà, è precisamente nella modalità di gestione di questa tensione – tra efficacia dell'intervento e garanzie di conoscibilità/controllo – che emergono differenze costituzionalmente rilevanti tra i modelli qui considerati.

⁷ *Provisions on the Management of Algorithmic Recommendations*, 2022; *Deep Synthesis Provisions e Generative AI Measures*, 2023.

⁸ Sul tema del ruolo degli algoritmi nei processi di Decision-Making cinesi, v. M. Wang, *Regulation of Algorithmic Decision-Making in China: Development, Problems and Implications*, 2024, 276 ss.

⁹ Sulla concorrenza tra modelli regolatori, v. A. Bradford, *Digital Empires: The Global Battle to Regulate Technology*, Oxford, 2023.

3. Le certificazioni vaccinali nell'Unione europea, negli Stati Uniti e in Cina

Nel contesto dell'emergenza pandemica, Unione europea, Stati Uniti e Repubblica Popolare Cinese hanno fatto ricorso – con intensità e finalità diverse – a dispositivi di certificazione digitale (variamente denominati “*Green Pass*” o “*Health Pass*”), capaci di trasformare uno status sanitario in un requisito operativo per l'accesso a mobilità, lavoro, servizi e spazi sociali. Ciò che si evidenzia, sul piano costituzionale, non è soltanto la compressione contingente di libertà e diritti, ma la tecnica con cui tale compressione viene resa eseguibile: standard di dati, regole di validità, procedure di verifica, responsabilità distribuite e, soprattutto, circuiti di controllo (democratici e indipendenti) più o meno robusti.

L'analisi comparata, come tra poco si vedrà, rivela come la stessa funzione (abilitare o negare l'accesso in base a informazioni sanitarie) viene incarnata in tre architetture di garanzia profondamente differenti, rendendo visibile il diverso equilibrio tra efficienza amministrativa, tutela dei diritti e qualità dell'*accountability*.

3.1. Il modello europeo tra interoperabilità tecnica e deficit di *accountability* preventiva

Nell'Unione europea¹⁰, l'*EU Digital COVID Certificate* (EUDCC) non va letto come un semplice strumento di verifica, bensì come un'autentica infrastruttura di governo della circolazione costruita in condizioni di emergenza.

La sua genesi è l'esito – non esente da fattori di criticità – di una regolamentazione accelerata¹¹, volta a consentire la libera circolazione nel mercato interno dopo i *lockdown*¹², mediante un meccanismo standardizzato di attestazione (vaccinazione, test, guarigione) in grado di operare come meccanismo autorizzatorio per l'accesso a spazi, servizi e mobilità.

La cornice giuridica è stata delineata dal regolamento (UE) 2021/953, che ha disciplinato l'*iter* di rilascio dei certificati di vaccinazione, test e guarigione ed è stato successivamente prorogato dal regolamento (UE) 2022/1034 fino al 30 giugno 2023¹³. Pur prevedendo espressamente che le limitazioni alla libertà di circolazione dovessero essere applicate «conformemente ai principi generali del diritto dell'Unione, segnatamente la proporzionalità e

¹⁰ F. Grasselli-V. Tevere, *Il “green pass” nello spazio europeo multilevel di libertà, sicurezza e giustizia. Riflessioni sull'eventuale introduzione dell'obbligatorietà vaccinale*, in *Freedom, Security & Justice: European Legal Studies*, 3, 2021, 243 ss.

¹¹ Commissione europea, *Relazione su un quadro per il rilascio, la verifica e l'accettazione di certificati interoperabili di vaccinazione, di test e di guarigione in relazione al COVID-19 per agevolare la libera circolazione delle persone durante la pandemica di Covid-19*, COM(2022) 753 final, Bruxelles, 22 dicembre 2022.

¹² Per la ricostruzione del quadro normativo dell'ordinamento europeo e italiano, v. C. Bertolino, *“Certificato verde Covid-19” tra libertà ed eguaglianza*, in *Federalismi.it*, 15, 2021, 2 ss.

¹³ Per un'ampia trattazione delle vicende che hanno condotto all'approvazione della normativa europea, v. T. Cerruti, *Il certificato digitale europeo tra luci e ombre*, in *EuroJus*, 2021, 17 ss.; D. Vese, *Certificare la salute? Il caso del Green Pass europeo*, in *ambientediritto.it*, 2022.

la non discriminazione»¹⁴, la fase di implementazione della certificazione ha rivelato un insufficiente quadro istruttorio trasparente e verificabile che consentisse un controllo preventivo robusto sulla necessità delle misure e sulle alternative meno intrusive¹⁵. Si questo specifico punto si colloca il nucleo di deficit di *accountability* preventiva, in quanto è mancato un dispositivo *ex ante* di giustificazione del bilanciamento tra salute, circolazione e altri diritti.

È significativo, in tale prospettiva, che l'EDPB e l'EDPS – nel Parere congiunto 04/2021 – abbiano segnalato la mancanza di una *Data Protection Impact Assessment* (DPIA)¹⁶ a corredo della proposta, proprio perché tale assenza ha indebolito la verificabilità (anche pubblica) dell'impatto delle misure e dell'effettività di soluzioni meno invasive. Difatti, com'è noto, una DPIA svolge, in termini di regolazione dei rischi tecnologici, la funzione di presidio della contestabilità preventiva, in quanto rende tracciabile la razionalità del progetto, consentendo di anticipare eventuali rischi di slittamento funzionale (*function creep*) e di discriminazione tra i vari interessi in gioco. L'assenza di una valutazione siffatta ha così spostato l'*accountability* dalla fase *ex ante* (dove dovrebbe pesare in misura maggiore) a controlli *in itinere* o *ex post*, con la conseguenza che la correzione degli errori è più lenta, più frammentata e politicamente più onerosa visto che l'infrastruttura è già entrata pienamente in funzione.

Il deficit di *accountability* preventiva è stato aggravato altresì dall'iter procedurale. Difatti, la scelta del Parlamento di attuare la *fast-track procedure* ha sacrificato gli spazi di verifica democratica in una fase delicata del *Law-Making*, dove la questione non è “se” il certificato funzioni, ma “quale” architettura dei diritti esso stabilizzi, a partire dalle soglie di validità, dalle possibili deroghe, dai meccanismi di motivazione e di contestazione del diniego fino alle garanzie contro effetti discriminatori su soggetti vulnerabili. In altre parole, il contesto emergenziale in cui gli organi comunitari si sono trovati ad operare ha privilegiato la capacità di implementazione rapida a scapito della costruzione di una legittimazione procedurale piena. Al fine di comprendere il punto, è utile precisare la catena di *accountability* che l'EUDCC ha prodotto. Da un lato, gli Stati membri hanno mantenuto la titolarità del rilascio e hanno definito gran parte delle scelte di *policy* (durata, requisiti, ambiti di applicazione) ma, dall'altro, la Commissione ha esercitato un ruolo centrale nel disegno dell'infrastruttura di fiducia, attraverso il *Gateway* europeo e la *governance* dell'interoperabilità¹⁷. Un ulteriore elemento di criticità ha riguardato la scelta della Commissione di riservarsi la possibilità di ridefinire il perimetro del trattamento dei dati mediante il

¹⁴ Regolamento (UE) 2021/953, considerando 6.

¹⁵ Si veda, in tema di mutamenti nella disciplina dei diritti fondamentali e sui nodi critici emersi con l'introduzione del *green pass* e obbligo vaccinale, A. Vidaschi-C. Graziani, *New Dynamics of the “Post-COVID-19 Era”: A Legal Conundrum*, in *German Law Journal*, 2023, 1613 ss.

¹⁶ Art. 35 GDPR.

¹⁷ Si veda la serrata analisi critica di O.J. Gstrein, *The EU Digital COVID Certificate: A Preliminary Data Protection Impact Assessment*, in *European Journal of Risk Regulation*, 2021, 373 ss.

ricorso ad atti delegati *ex art.* 290 TFUE¹⁸, così da operare aggiustamenti rapidi, anche in relazione all'evoluzione della situazione epidemiologica. Ciò ha accresciuto la capacità di adattamento tecnico-scientifico ma, al contempo, ha accentuato la dipendenza del sistema da un circuito decisionale meno "parlamentarizzato" rispetto alla fase legislativa ordinaria.

Tale filiera di responsabilità ha prodotto un effetto tipico della regolazione per standard, rendendo scalabile e portabile la decisione politica incorporandola in un'infrastruttura tecnica, con il rischio che la trasparenza si concentri sul "come" (protocollo, sicurezza, interoperabilità) più che sul "perché" (giustificazione delle soglie e dei *trade-off*). Non a caso, EDPB/EDPS hanno altresì osservato che modifiche ai *data fields* tramite atti delegati possono incidere sul profilo di rischio e richiedere una nuova valutazione, segnalando, ancora una volta, la centralità del controllo preventivo. Sul piano tecnico, il modello europeo ha adottato un'infrastruttura a chiave pubblica (PKI) con specifiche tecniche accessibili e componenti a codice sorgente aperto¹⁹, privilegiando l'interoperabilità del sistema e consentendo audit di sicurezza e verifiche indipendenti sul funzionamento della firma digitale e della verifica *offline*. Tale scelta non solo ha consentito a studiosi e autorità garanti di condurre *audit* tecnici del codice, ma ha altresì favorito il riconoscimento reciproco dei certificati tra i 27 Stati membri, nonché l'integrazione – nel circuito europeo²⁰ – di oltre sessanta paesi terzi attraverso procedure di "*onboarding*"²¹. Sotto altro profilo, l'architettura distribuita, basata su verifica locale mediante chiavi crittografate pubbliche senza trasmissione dei dati sanitari a livello centrale, ha limitato i rischi di sorveglianza di massa.

Tuttavia, tale elevata trasparenza ha riguardato solamente l'implementazione tecnica e non anche i criteri di *policy*, quali le soglie di validità, logiche delle esenzioni e degli accessi condizionati, le ragioni del rigetto di opzioni meno invasive e, soprattutto, l'assetto complessivo di responsabilità e rimedi. Le ricadute di tali scelte hanno prodotto una trasparenza di tipo asimmetrico, con alti indici di accessibilità e spiegabilità per chi possedeva competenza tecniche, ma più debole per i cittadini che subivano decisioni in termini di diniego di accesso in determinati luoghi e necessitavano, nell'immediato, di motivazioni comprensibili e di una chiara individuazione dei soggetti responsabili. È chiaro che in un sistema che ha operato come infrastruttura di accesso, l'*accountability ex post* non è risultata strutturalmente inadeguata in quanto non supportata da strumenti di contestazione immediata e da una tracciabilità della decisione²².

In termini di diffusione, la Commissione ha documentato l'emissione di

¹⁸ Difatti, il regolamento (UE) 2021/953, all'art. 12, delega alla Commissione il potere di integrare o modificare elementi non essenziali, inclusi aspetti tecnici e i set di dati contenuti nei certificati.

¹⁹ *EU Digital COVID Certificate*, repository GitHub.

²⁰ *Ibid*

²¹ Commissione europea, *Certificato COVID digitale dell'UE*.

²² Sul concetto di *accountability* come relazione complessa e articolata tra "*answerability*" ed "*enforcement*", v. M. Bovens, *Analysing and Assessing Accountability: A Conceptual Framework*, in *European Law Journal*, 13, 2007, 447 ss.; Id., *Two Concepts of Accountability as a Virtue and as a Mechanism*, in *West European Politics*, 33, 5, 2010, 946 ss.

oltre 2,2 miliardi di certificati nel periodo di operatività del sistema (giugno 2021-giugno 2023)²³, configurando il “green pass” europeo come il più vasto sistema di certificazione digitale interoperabile mai implementato a livello sovranazionale. Il suddetto dato certifica senz’altro il successo operativo in termini di diffusione e funzionamento tecnico, ma non risolve l’interrogativo sulla qualità dello scrutinio preventivo dei rischi strutturali per i diritti fondamentali. La redazione di *report* periodici²⁴, il coinvolgimento dei Garanti nazionali e la disponibilità di rimedi giurisdizionali hanno solo parzialmente compensato l’assenza di una valutazione indipendente e pubblica dei rischi nella fase costitutiva del sistema.

L’esperienza europea restituisce, dunque, un modello ambivalente, con alta trasparenza algoritmica a fronte di opacità decisionale nella fase costitutiva, mentre, sul versante dell’*accountability*, esso si è contraddistinto per frammentazione delle responsabilità, prevalenza di controlli *ex post* e rimedi solo parzialmente efficaci. Dunque, l’urgenza emergenziale ha compreso le garanzie procedurali, generando un *deficit* di trasparenza *ab initio*, difficilmente colmabile a posteriori. Tale assunto è supportato dall’analisi del funzionamento del sistema EUDCC a livello nazionale²⁵, in relazione alla quale si è evidenziato – come di seguito si vedrà – significative divergenze nelle modalità di utilizzo e nelle implicazioni sul piano della trasparenza algoritmica e *accountability*²⁶.

L’Italia ha implementato il sistema più estensivo e pervasivo tra gli Stati membri di certificazione digitale, attraverso una stratificazione normativa complessa, avviata con il decreto-legge 22 aprile 2021, n. 52²⁷ e successivamente estesa mediante i decreti-legge nn. 105²⁸, 127²⁹ e 172³⁰ del 2021³¹. Operativamente, il “green pass” italiano ha operato secondo tre livelli di

²³ Commissione europea, *Relazione della Commissione al Parlamento europeo e al Consiglio su un quadro per il rilascio, la verifica e l'accettazione di certificati interoperabili di vaccinazione, di test e di guarigione in relazione al COVID-19 per agevolare la libera circolazione delle persone durante la pandemia di COVID-19*, COM(2022) 753 final, Bruxelles, 22 dicembre 2022.

²⁴ *Report from the Commission to the European Parliament and the Council*, COM(2022) 123 final, Brussels, 15 marzo 2022.

²⁵ Per un’analisi ad ampio spettro delle ricadute dei certificati digitali su 9 Paesi europei, v. L. Marelli-B. Zimmermann-H. Machado, *et al.*, *COVID-19 Health Certificates: analysis of policy implementation, framings, and controversies in nine European countries*, in *Science and Public Policy*, 2025.

²⁶ Cfr. C. Bertolino, “Certificato verde Covid-19” tra libertà ed eguaglianza, cit., 2 ss.

²⁷ Il Decreto in parola introdusse la “certificazione verde COVID-19” per spostamenti tra regioni di diverso colore e per partecipazione a eventi.

²⁸ Il d.l. 105/2021 ampliò l’uso del *green pass* a ristoranti al chiuso, cinema, eventi pubblici e fiere.

²⁹ Il d.l. 127/2021 impose il *green pass* a tutto il personale della scuola, università e sanità e, a partire dal 15 ottobre 2021, a tutti i lavoratori pubblici e privati.

³⁰ Il d.l. 172/2021 introdusse il rafforzamento del sistema con il cd “super green pass”, rilasciato solo a vaccinati e guariti, non più ottenibile con il solo tampone, per attività ricreative e di socialità.

³¹ Per una ricostruzione dell’iter normativo italiano e una disamina sulla portata costituzionale del “green pass” italiano, v. L. Brunetti, *La Costituzione può obbligare, ma non costringere: sulla incostituzionalità degli attuali divieti come mezzo per incentivare la vaccinazione anti COVID-19 (cd green pass)*, in *Forum di Quaderni costituzionali*, 3, 2021.

obbligatorietà: inizialmente richiesto per accesso a eventi, ristorazione al chiuso e strutture ricettive (nella formula di test antigenico negativo, guarigione o vaccinazione), successivamente è stato esteso a tutto il personale sanitario, scolastico e universitario (c.d. “*green pass base*”); infine, a partire da dicembre 2021, reso obbligatorio per attività ludico-ricreative (ristoranti, spettacoli, sport di squadra, palestre e discoteche con il c.d. “*super green pass*”, ottenibile solo con vaccinazione o guarigione)³². Tale estensione ha generato un abnorme contenzioso costituzionale. La Corte costituzionale non ha mai dichiarato l’illegittimità del sistema, riconoscendone la legittimità come strumento di tutela della salute pubblica, nel quadro di una lettura centrata su proporzionalità e temporaneità delle misure³³.

Sul versante della trasparenza, l’uso dello standard europeo EUDCC (firma digitale, verifica offline), codice sorgente aperto e minimizzazione dei dati ha reso il funzionamento controllabile e tendenzialmente interoperabile. Tuttavia, la definizione delle soglie e dei criteri di validità è avvenuta in larga misura tramite una normativa stratificata e complessa, con frequenti rinvii a specifiche tecniche. Ne è derivata una trasparenza sbilanciata, con un grado elevato sul versante del codice e del protocollo, più debole sulle scelte di *policy*³⁴.

Anche sul piano dell’*accountability*, l’obbligo di verifica del “green pass” è stato spostato su esercenti, datori di lavoro e organizzatori, con un processo sanzionatorio a valle, producendo frammentazione delle responsabilità e asimmetrie informative tra controllori e controllati³⁵. Anche sul piano

³² Per una ricostruzione del complesso iter e del controverso rapporto tra green pass e obbligo vaccinale, v. A.M. Citrigno, *Obbligo del green pass e obbligo vaccinale: le questioni aperte*, in *dirittifondamentali.it*, 1, 2022, 2 ss.

³³ Nelle fondamentali sent. n. 14 e 15 del 2023, la Corte costituzionale ha ricondotto l’obbligo vaccinale anti-Covid entro il perimetro dell’art. 32 Cost., ribadendo, nella sent. n. 15/2023, che l’obbligo tutela non solo il lavoratore ma la comunità – specie i soggetti fragili – e serve a preservare l’operatività dei servizi, pertanto, la sospensione del lavoratore non vaccinato costituisce misura di sanità pubblica, ragionevole e proporzionata, destinata a cessare con l’adempimento. Con la sent. n. 14/2021, la Corte chiarisce poi che l’obbligo si accompagna a un consenso informato che, in tale contesto, ha funzione informativa e di triage clinico e che l’eventuale rischio di eventi avversi non invalida la legge ma integra, semmai, il titolo all’indennizzo. Più recentemente, con la sent. n. 199/2025, la Corte costituzionale ha dichiarato che il “Green Pass” e l’obbligo vaccinale per tutti i soggetti “over 50”, con relativa sospensione della retribuzione per i non vaccinati, sono da ritenersi misure conformi al dettato costituzionale, in quanto, durante un’emergenza sanitaria, la salute pubblica prevale sulla scelta individuale di non adempiere agli obblighi di sicurezza sul lavoro. Per ulteriori approfondimenti, cfr. C. Iannello, *La sentenza n. 14/2023 della Corte costituzionale: l’obbligo vaccinale è legittimo solo se serve a prevenire il contagio*, in *Osservatorio AIC*, 4, 2023, 80 ss.; G. Cerrina Feroni, *Obblighi vaccinali, conseguenze del mancato assolvimento e Costituzione. Una lettura critica delle sentenze della Corte costituzionale n. 14 e 15 del 2023*, in *dirittifondamentali.it*, 2, 2023, 245 ss.; A. Iannuzzi, *Vaccinazioni obbligatorie, dati scientifici e consenso informato: a margine delle sentenze n. 14 e 15 del 2023 della Corte costituzionale*, in *Nomos*, 2, 2023, 2 ss.; G. Vecchio, *Riflessioni sulla sospensione della retribuzione a favore dei lavoratori obiettori dopo le sentenze della Corte costituzionale n. 14 e 15 del 2023*, in *dirittifondamentali.it*, 1, 2023, 425 ss.

³⁴ Su questi temi, v. C. Bertolino, “Certificato verde Covid-19” tra libertà ed eguaglianza, cit., 22.

³⁵ Per ulteriori approfondimenti sulla portata del “green pass”, v. R. Berti-F. Zumerle, *Obbligo Green pass in Italia, le novità sul filo della Costituzione*, in *agendadigitale.eu*, 26 marzo 2022.

dei rimedi, i processi di contestazione erano spesso lenti e distribuiti tra innumerevoli soggetti (ASL, piattaforme di emissione, datori di lavoro), rendendo la tutela giurisdizionale complessa e onerosa, anche in ragione del contesto tecnico complesso e in continua evoluzione³⁶.

Anche la Francia ha adottato un sistema analogo a quello italiano³⁷ attraverso la “*loi relative à la gestion de la crise sanitaire*”, del 5 agosto 2021, che ha introdotto il “*pass sanitaire*” obbligatorio per accesso a ristoranti, bar, trasporti di lunga percorrenza, centri commerciali e strutture sanitarie³⁸, poi evoluto in “*pass vaccinal*” con la legge del 22 gennaio 2022. Il *Conseil constitutionnel*³⁹, ha validato l’impianto dell’intero sistema, introducendo però limiti e riserve, con esiti di non conformità parziale su specifici profili⁴⁰.

Sul piano tecnico, anche la Francia ha adottato lo standard europeo “EU-DCC” con architettura di firma digitale e verifica offline tramite “*TousAntiCovid Verif*”, oggetto di molteplici pareri da parte del CNIL che ha richiesto garanzie rafforzate e valutazioni d’impatto⁴¹. Anche nel caso francese, è emersa una forte asimmetria tra un alto grado di trasparenza tecnico-scientifica dell’algoritmo di verifica, ma una accentuata opacità sui criteri di *policy* – durata, esenzioni, soglie – che abilitavano o escludevano in tempo reale⁴². quanto all’*accountability*, anche in Francia l’onere dei controlli è caduto su esercenti, gestori e datori di lavoro, con un regime sanzionatorio a valle.

La Germania, per contro, ha adottato un approccio meno invasivo, limitando l’uso del certificato principalmente ai viaggi internazionali e all’accesso a eventi specifici, anche si vi è stata una fase – dal novembre 2021 al marzo 2022 – in cui l’uso del *pass* è stato esteso ai luoghi di lavoro su base federale, con obblighi di verifica in capo ai datori di lavoro e successivo obbligo di cancellazione dei dati alla scadenza della norma. Operativamente, anche il sistema tedesco ha adottato lo standard “EUDCC”, su base decentralizzata, con i certificati che venivano memorizzati sul dispositivo e verificati *offline* tramite firma crittografata, minimizzando la tracciabilità⁴³. Il *Bundestag* ha approvato modifiche all’*Infektionsschutzgesetz* che hanno per-

³⁶ In relazione alla controversa portata del “*green pass*” sui luoghi di lavoro, v. L. Bartolucci, *Contributo allo studio dei profili costituzionali del “green pass”: trattamento sanitario obbligatorio “indiretto” e “di fatto”?*, in *Federalismi.it*, 23, 2022, 3 ss.

³⁷ Per una comparazione tra il sistema italiano e quello francese, cfr. G. Stegher, *Gli stati costituzionali di diritto alla prova del “green pass”: Italia e Francia a confronto*, in *DPCE online*, 2022, 662 ss.

³⁸ Per un approfondimento sulla portata del “*pass sanitaire*”, v. S. Bandelier-C. Dressel-C. Lanier et al., *Retour sur le pass vaccinal: libre disposition de son corps et consentement aux soins*, in *La Revue des Droits de l’Homme*, 2022.

³⁹ Per un ampio e approfondito commento della sentenza, v. G. Grasso, *Il passe sanitaire e i diritti e le libertà costituzionalmente garantiti nella decisione n. 2021-824 del 5 agosto 2021 del Conseil constitutionnel*, in *DPCE online*, 3, 2021, 2819 ss.

⁴⁰ Cfr., in merito, V. Sizaire, *Un passeport pour a coercition?*, in *La Revue des Droits de l’Homme*, 2022.

⁴¹ I pareri del CNIL sono consultabili online.

⁴² *Approfondimenti* sono consultabili online sul sito del CNIL.

⁴³ A.K. Mangold, *Germany and Covid-19: A Most Eventful Year*, in *verfassungsblog.de*, 7 aprile 2021.

messo l'utilizzo del certificato ma con garanzie procedurali più robuste e *sunset clauses* automatiche⁴⁴.

Tuttavia, anche in questo caso, la verifica *offline* e l'assenza di *logging* sistematico se, per un verso, hanno tutelato la *privacy*, per altro verso, ha limitato eventuali audit *ex post*, per controllare la persistenza di *bias* algoritmici o disparità nella fase applicativa.

Rispetto a Francia e Italia, l'approccio tedesco ha privilegiato proporzionalità, temporalità e verificabilità tecnica, traducendosi, in termini di *governance* degli algoritmi, in trasparenza *ex ante* sul funzionamento dell'algoritmo e *accountability* rafforzata dalla previsione di *sunset clauses*, pur con il prezzo di una auditabilità *ex post* meno penetrante per effetto della scelta di verifiche *offline* e minimizzazione dei *log*.

Altri Stati membri hanno sviluppato approcci intermedi. La Spagna ha utilizzato il certificato principalmente per viaggi e accesso a eventi ma non per luoghi di lavoro ordinari⁴⁵; i Paesi Bassi hanno introdotto il sistema ma lo hanno rapidamente sospeso dopo proteste pubbliche significative⁴⁶. Infine, la Danimarca⁴⁷ è stata tra i primi paesi ad abolire completamente il sistema già nel settembre 2021⁴⁸.

Le Autorità Garanti nazionali hanno svolto un ruolo cruciale nel garantire il rispetto dei principi del GDPR nell'attuazione dell'EUDCC. Il Garante italiano ha subordinato il parere favorevole sul DPCM attuativo della piattaforma nazionale DGC a condizioni stringenti, richiedendo che la verifica avvenisse mediante QR Code senza conservazione dei dati e accesso limitato alle sole informazioni strettamente necessarie (validità/invalidità del certificato, senza dettagli sul tipo di certificazione)⁴⁹. La CNIL francese ha imposto limitazioni analoghe, vietando la conservazione dei dati di verifica e richiedendo trasparenza sugli algoritmi utilizzati⁵⁰. Il Bundesamt für Sicherheit in der Informationstechnik (BSI) tedesco ha condotto *audit* di sicurezza approfonditi sul sistema “Corona-Warn-App” sin dal suo sviluppo, supportando la progettazione dell'app in termini di sicurezza IT, esaminando costantemente le versioni del software e le infrastrutture correlate, nonché fornendo consulenza sul concetto di sicurezza da adottare⁵¹.

Il quadro europeo consegna un modello tecnicamente trasparente e fun-

⁴⁴ Cfr. S. Jürgensen-F. Orlowski, *Critique and Crisis: The German Struggle with Pandemic Control Measures and the State of Emergency*, in *verfassungsblog.de*, 19 aprile 2020.

⁴⁵ Cfr. A. Vedaschi, *COVID-19 and Emergency Powers in Western European Democracies: Trends and Issues*, in *verfassungsblog.de*, 5 maggio 2021.

⁴⁶ Cfr. M. Julicher-M. Vetzo, *COVID-19 in the Netherlands: of Chancing Tides and Constitutional Constants*, in *verfassungsblog.de*, 22 aprile 2021.

⁴⁷ Cfr. K. Cedervall Lauta, *Something is Forgotten in the State of Denmark: Denmark's Response to the COVID-19 Pandemic*, in *verfassungsblog.de*, 4 maggio 2020.

⁴⁸ Danish Health Authority, *In the future, your corona vaccination passport will expire sooner*, 9 dicembre 2021.

⁴⁹ Garante per la protezione dei dati personali, *Parere sul DPCM di attuazione della piattaforma nazionale DGC per l'emissione, il rilascio e la verifica del Green Pass*, 9 giugno 2021.

⁵⁰ Commission Nationale Informatique & Libertés, *Délibération n. 2021-054 du 12 mai 2021 portant avis sur le projet de mise en place d'un passe sanitaire conditionnant l'accès à certains lieux, événements ou établissements impliquant de grands rassemblements de personnes*, 12 maggio 2021.

⁵¹ Bundesamt für Sicherheit in der Informationstechnik, *Die Corona-Warn-App*, 2021.

zionalmente efficace, ma afflitto da una criticità strutturale, in relazione alla difficoltà di costruire un’accountability preventiva all’altezza di un’infrastruttura destinata a diventare standard. La protezione dei dati ai controlli *in itinere* ed *ex post* – pareri delle Autorità Garanti, linee guida EDPB, rimedi giurisdizionali – che hanno colmato solo parzialmente il *deficit* di *accountability* genetica. Dunque, l’esperienza europea evidenzia come l’urgenza possa comprimere l’istruttoria pubblica iniziale, i controlli *ex post* – per quanto numerosi – faticano a ricostruire quella legittimazione procedurale che avrebbe dovuto presidiare il sistema prima della sua generalizzazione.

3.2. Il modello statunitense tra frammentazione federale e *governance* privatistica

Diversamente dall’esperienza europea, gli Stati Uniti non hanno implementato un “passaporto vaccinale” federale né un’infrastruttura nazionale paragonabile all’EUDCC. Tale scelta non è stata meramente casuale, bensì il precipitato di due diversi fattori. Da un lato, l’assetto federale della sanità pubblica e, dall’altro, l’elevata sensibilità politico-costituzionale verso strumenti percepiti come potenziali elementi di violazione della *privacy*. In tale cornice, la posizione dell’amministrazione federale è stata improntata alla massima prudenza, lasciando spazio a soluzioni statali e, soprattutto, di mercato.

Il vuoto di una cornice federale è stato colmato, anzitutto, da un processo di standardizzazione non legislativa, promosso da consorzi e attori industriali. La “*Vaccination Credential Initiative*” (VCI)⁵², ha avuto un ruolo catalizzatore⁵³, impegnando organizzazioni partecipanti a implementare il framework delle “*SMART Health Cards*” e a rilasciare credenziali in forme compatibili, secondo guide di implementazione di matrice tecnica.

L’elemento di maggiore interesse è relativo al fatto che la “uniformità” viene conseguita non attraverso una fonte normativa, di portata generale, ma tramite un insieme di regole consortili, specifiche tecniche e pratiche di mercato. Operativamente, le “*SMART Health Cards*” si sono basate su standard sanitari interoperabili e su firme crittografate (tipicamente in codici QR), consentendo verifica e portabilità senza imporre, tuttavia, un registro federale unico dei vaccinati.

⁵² La *Vaccination Credential Initiative* è nata da una *partnership* pubblico-privata lanciata nel gennaio 2021 cui hanno partecipato Mayo Clinic, Microsoft, Oracle, Salesforce e numerosi altri attori sanitari e tecnologici. Cfr., *Vaccination Credential Initiative, About VCI*.

⁵³ Basti pensare che alcuni Stati, come New York, hanno sviluppato sistemi proprietari basati su tecnologia IBM (New York State, *Excelsior Pass*), mentre altri, come il caso dello Stato della California, hanno adottato le *SMART Health Cards* attraverso il *Digital COVID-19 Vaccine Record* (California Department of Public Health, *Digital Covid-19 Vaccine Record*) e con molti altri Stati che hanno vietato l’uso di “*vaccine passports*”, tra cui Florida, Texas, Alabama, Arizona e Arkansas (National Conference of State Legislatures, *COVID-19 Legislation Archived Database*, 2022-2023). Tale eterogeneità ha creato un *patchwork* giuridico, che ha sollevato non poche questioni di costituzionalità sulla *Privileges and Immunities Clause* (art. IV, sez. 2) e sulla *Commerce Clause*. Cfr., per un’analisi più approfondita, M.M. Mello-D. J. Opel-R. M Benjamin *et al.*, *Effectiveness of vaccination mandates in improving uptake of COVID-19 vaccines in the USA*, in *Lancet*, 13 agosto 2022.

Tale architettura, che in termini di efficienza è stata decisiva, ha spostato tuttavia il baricentro della decisione su un circuito prevalentemente privatistico, in cui la trasparenza è anzitutto quella del protocollo e delle specifiche, non quella di una procedura democratica di bilanciamento paragonabile a un *Law-Making* pubblico.

In assenza di un quadro federale, alcuni Stati hanno optato per soluzioni proprie. Tra questi, la California ha implementato un “*Digital COVID-19 Vaccine Record*” verificabile tramite lettori compatibili con “*SMART Health Cards*”⁵⁴, rendendo pubblico un *repository* tecnico relativo al QR e pubblicando su GitHub il codice del portale per l’emissione di tali credenziali⁵⁵. Lo Stato di New York ha sviluppato l’“*Excelsior Pass*”, una credenziale digitale per vaccinazioni/test COVID, integrata anch’essa nel framework “*SMART Health Cards*” tramite a “*Vaccination Credential Initiative*”. Nel 2021, lo Stato ha pubblicato l’“*Excelsior Pass Blueprint*”, un documento ufficiale che illustra l’architettura, i componenti, gli aspetti di privacy e interoperabilità, ma non il codice sorgente completo né i dettagli algoritmici⁵⁶.

Lo snodo giurisprudenziale che chiarisce la grammatica costituzionale dell’esperienza pandemica negli USA è offerto da due decisioni della Corte Suprema del gennaio 2022. In “*NFIB v. Department of Labor (OSHA)*”, i giudici hanno confermato che i “*police powers*” in materia di salute pubblica risiedevano principalmente negli Stati, bloccando il mandato vaccinale di OSHA per le grandi imprese, in quanto ritenuto eccedente l’ambito competenziale dell’agenzia in materia di sicurezza sul lavoro⁵⁷. Contestualmente, in “*Biden v. Missouri*”⁵⁸, ha invece confermato la legittimità del mandato vaccinale del *Centers for Medicare & Medicaid Services* (CMS) per operatori sanitari di strutture sanitarie partecipanti a “*Medicare/Medicaid*”, evidenziando il nesso tra condizioni di partecipazione ai programmi federali e tutela dei pazienti. Tali pronunce, pur non riguardando direttamente il *pass* vaccinale, hanno comunque inciso sull’ordinamento, delimitando ciò che può essere imposto a livello federale tramite agenzie e, conseguentemente, rafforzando il ruolo degli Stati e degli attori privati nello spazio in cui le credenziali digitali diventano operative.

⁵⁴ Cfr. spec.smarthealth.cards.

⁵⁵ Per ulteriori informazioni, cfr. [Repository Github](https://github.com), in github.com.

⁵⁶ Per ulteriori approfondimenti tecnici, cfr. *New York State Excelsior Pass – Blueprint*, in ny.gov.

⁵⁷ È interessante notare come la Corte abbia fatto leva sulla “*major questions doctrine*” – canone interpretativo secondo cui un’agenzia amministrativa non può rivendicare il potere di decidere questioni di «*vast economic and political significance*» in assenza di una chiara autorizzazione legislativa – per sospendere (e di fatto paralizzare) il mandato vaccinale adottato da OSHA. Come ha sottolineato il giudice Gorsuch nella *concurring opinion*, permettere a OSHA di affrontare rischi sanitari generali attraverso la normativa sulla sicurezza nei luoghi di lavoro avrebbe conferito all’agenzia un potere «*unprecedented*» di incidere sulla vita di «*84 million Americans*», trasformando la disciplina della sicurezza sul lavoro in uno strumento di regolazione sanitaria generale, ossia un salto di competenza che, secondo l’opinione concorrente, avrebbe richiesto una chiara base legislativa. Supreme Court, *National Federation of Independent Business v. Department of Labor*, 142 S. Ct. 661, 2022.

⁵⁸ Supreme Court, *Biden v. Missouri*, 142 S. Ct. 647 (2022).

Dunque, il modello privatistico di credenziali sanitarie digitali ha spostato il *locus* del potere dal diritto pubblico verso regole organizzative e contrattuali, spesso implementate da soggetti che non sono “*state actors*” in senso stretto. Da ciò ne è derivato un *vulnus* evidente in termini di trasparenza e *accountability*, atteso che una parte rilevante delle garanzie pubblicistiche (*due process* procedimentale, obblighi di motivazione, sindacato tipico delle decisioni amministrative) si attiva con difficoltà quando la decisione “che esclude” è formalmente privata, anche se funzionalmente agganciata a finalità pubbliche⁵⁹.

Anche se la trasparenza tecnica è stata garantita da un adeguato livello di interoperabilità dei vari sistemi di certificazioni, la trasparenza sostanziale, ossia la comprensibilità delle regole, è apparsa decisamente più fragile. Le soglie di validità, le esenzioni, le alternative accettate sono dipese da un mosaico di ordinanze statali e locali, *policy* aziendali e clausole contrattuali. Sicché, anche quando le credenziali erano tecnicamente chiare e verificabili, le ragioni per cui esse potevano essere ritenute valide o meno in un determinato luogo sono rimaste opache o mutevoli.

In tale contesto, è chiaro che anche sul fronte dell’*accountability* il risultato è stato contrastante, in quanto, a fronte di una forte innovazione tecnico-organizzativa garantita dalle piattaforme, la responsabilizzazione pubblica è risultata asimmetrica e frammentata, compromettendo la trasparenza sostanziale delle scelte, l’effettività dei rimedi per i singoli e, in ultima analisi, il controllo democratico sugli standard algoritmici che regolano l’accesso a diritti e servizi essenziali, soprattutto nella drammatica fase pandemica.

3.3. Il modello cinese: dinamismo algoritmico e opacità sistemica

Il sistema cinese degli “*Health Codes*” si colloca su un piano strutturalmente diverso rispetto ai modelli europeo e statunitense. Non si tratta soltanto di “certificare” uno status sanitario, ma di tradurre in tempo quasi reale un giudizio di rischio in una condizione di accesso alla vita sociale. Mentre il “greenpass” europeo era ancorato a presupposti relativamente stabili e comprensibili (vaccinazione, test, guarigione), l’esperienza cinese ha assunto la forma di un dispositivo di governo della mobilità, in grado di comprimere o riattivare la libertà di circolazione dei cittadini sulla base di segnali di rischio non pienamente intellegibili per i destinatari⁶⁰.

Dal punto di vista operativo, gli “*Health Codes*”, integrati nelle piattaforme *WeChat* e *Alipay* che permeano capillarmente la vita quotidiana cinese, sono diventati, in breve tempo, un passaggio obbligato per l’accesso a spazi e servizi: trasporti, uffici, luoghi di lavoro, centri commerciali, strutture sanitarie e, in molte città, persino l’ingresso nei complessi residenziali. Il codice, tipicamente espresso tramite QR e colori – verde (libera circolazio-

⁵⁹ Cfr. P.A. Diller, *Training a Public Accommodations Lens on Vaccine Passports*, in *Lewis & Clark Law Review*, 523, 2023, 525 ss.

⁶⁰ Per un’ampia panoramica sui caratteri della governance digitale cinese nell’emergenza pandemica, cfr. F. Liang, *Covid-19 and Health Code: How Digital Platform Tackle the Pandemic in China*, in *Social Media + Society*, 2020.

ne), giallo (quarantena di 7 giorni), rosso (quarantena di 14 giorni) – non si limitava a certificare uno status, ma abilitava o bloccava concretamente la circolazione e l'accesso, rendendo gli effetti immediati. Il sistema era completato dalla “*Communication Itinerary Card*”, basata sui dati degli operatori di telecomunicazioni, che tracciava automaticamente tutte le città visitate negli ultimi 14 giorni dove lo *smartphone* aveva sostato per più di 4 ore, alimentando gli algoritmi di rischio con dati oggettivi di mobilità⁶¹.

Il punto costituzionalmente sensibile è che, diversamente da un'attestazione statica, l'Health Code cinese si è presentato spesso come una classificazione algoritmica del rischio, che considerava la cronologia degli spostamenti, i contatti con persone risultate positive al coronavirus o ad alto rischio, i risultati di un test antigenico, la permanenza in aree classificate ad alto rischio e, dal maggio 2021, lo stato vaccinale visualizzato attraverso indicatori come bordi dorati o mascotte simboliche⁶². Diversamente dal modello europeo, dove lo statuto certificativo rimaneva stabile salvo eventi determinati (scadenza test, nuova vaccinazione), il sistema cinese operava mediante valutazione continua e retroattiva. Il codice poteva mutare automaticamente, senza intervento dell'utente né notifica preventiva. In altre parole, un cittadino poteva scoprire di avere improvvisamente il codice rosso perché l'algoritmo aveva rilevato retroattivamente un contatto giorni prima o un transito in zone successivamente classificate ad alto rischio⁶³. Dunque, sul piano giuridico, la trasparenza è stata sostituita dalla pervasività infrastrutturale dell'“*Health Code*”, in cui il punteggio di rischio si fondava su una fusione massiva di dati – geolocalizzazione, trasporti, telecomunicazioni, pagamenti elettronici, videosorveglianza – che abilitava una verifica capillare delle prescrizioni ma, al contempo, rendeva opachi i criteri di classificazione e i possibili errori di esito. In breve, ciò garantiva altissima controllabilità ma bassa spiegabilità. In tale contesto, l'inasprimento sanzionatorio per le violazioni ha poi contribuito a rafforzare l'efficacia dello strumento più che la qualità informativa verso i cittadini⁶⁴. Tale dinamicità configura una differenza costituzionalmente rilevante: mentre il “*greenpass*” certificava uno *status* verificabile *ex ante* dal titolare, l'“*Health Code*” costituiva, appunto, una valutazione di rischio opaca, modificabile unilateralmente dall'algoritmo e sottratta a contestazione⁶⁵. Tale architettura realizza ciò che la dottrina definisce “*hypernudge*”⁶⁶, ossia un

⁶¹ *Ibid.*

⁶² Per una più approfondita disamina, anche tecnica, cfr. H. Yu, *Living in the era of codes: a reflection on China's health code system*, 2022.

⁶³ Sul funzionamento del sistema di “*Health Code*”, cfr. W. Cong, *From Pandemic Control to Data-Driven Governance: The Case of China's Health Code*, in *Frontiers in Political Science*, 14 aprile 2021.

⁶⁴ Cfr. X. Gao-J. Yu, *Public governance mechanism in the prevention and control of the Covid-19: information, decision-making and execution*, in *Journal of Chinese Governance*, 14 aprile 2020; J. Zhang-R. Zhang, *COVID-19 in China: Power, Transparency and Governance in Public Health Crisis*, in *Healthcare*, 2020.

⁶⁵ Cfr. X. Zhang, *Decoding China's COVID-19 Health Apps: The Legal Challenges*, in *Healthcare*, 2022.

⁶⁶ La definizione rinvia al lavoro pionieristico di K. Yeung, che ha introdotto il concetto di “*hypernudge*” nel contesto dell'architettura delle scelte digitali alimentata dall'uso di Big

sistema di condizionamento comportamentale pervasivo e personalizzato che opera al di sotto della soglia di consapevolezza individuale, erodendo l'autonomia decisionale senza ricorrere a divieti formali espliciti.

La base legale degli “*Health Codes*” è costituita da un *corpus* normativo particolarmente articolato – *Cybersecurity Law* del 2017, *Data Security Law* del 2021, *Personal Information Protection Law* del 2021⁶⁷ – in tema di tutela dei dati personali e di sicurezza dei dati, formalmente comparabili al GDPR, in quanto introduce principi generali su consenso, finalità, protezione e trasparenza nel trattamento dei dati personali⁶⁸. Nonostante tale quadro di norme e principi, la gestione operativa degli “*Health Codes*” è stata tuttavia regolata da una fitta produzione di direttive e circolari tecniche emanate dalla *National Health Commission* e del *Joint Prevention and Control Mechanism*. Tali atti amministrativi, spesso emergenziali e poco trasparenti, hanno di fatto largamente derogato l'applicazione sostanziale delle garanzie formali previste dalla legge, rendendo difficile la contestazione giurisdizionale da parte dei cittadini. Difatti, la supervisione effettiva da parte dei tribunali e di altre istituzioni indipendenti è stata limitata in virtù della prevalenza delle esigenze di salute pubblica e del controllo amministrativo centralizzato⁶⁹. Per tale ragione, la letteratura ha messo in luce non poche lacune su consenso, trasparenza, necessità e proporzionalità e rimedi, soprattutto nella fase di transizione dall'emergenza alla normalizzazione⁷⁰.

Sul piano vaccinale, la Cina non ha mai implementato un obbligo formale per specifiche categorie professionali, intervenendo anche per frenare iniziative locali – come quella avviata a Pechino nel luglio 2022 – che tendevano a trasformarla in requisito generalizzato di accesso ai luoghi pubblici⁷¹. Tuttavia, non sono mancate forme di indebita pressione. Difatti, in molte aree del Paese, l'adesione alla campagna vaccinale è stata sostenuta da meccanismi indiretti di incentivi e coercizione sociale-amministrativa e, in taluni casi, da condizionamenti pratici nell'accesso a servizi o nella vita quotidiana⁷².

Data e algoritmi predittivi. Yeung descrive l’“*hypernudge*” come una sofisticazione della tradizionale architettura delle scelte che, grazie, ai Big Data e alla profilazione dinamica, personalizza e reindirizza in tempo reale le opzioni disponibili all'utente, influenzando profondamente i processi decisionali. Cfr. K. Yeung, “*Hypernudge*”: *Big Data as a Mode of Regulation by Design*, in *Information, Communication & Society*, 20, 1, 2017, 118 ss.

⁶⁷ Cfr. R. Demaria, *La data regulation in Cina tra politica interna e proiezione esterna*, in *OrizzonteCina*, 13, 2-3, 2022, 109 ss.

⁶⁸ In tema, cfr. la fondamentale opera di R. Creemers, *China's Emerging Data Protection Framework*, 2021.

⁶⁹ Cfr. H. Dorwart-G. Zanfir-Fortuna-C. Giro, *China's new comprehensive data protection law: context, stated objectives, key provisions*, in *Future of Privacy Forum*, 20 agosto 2021.

⁷⁰ Cfr. X. Zhang, *Decoding China's COVID-19 Health Apps: The Legal Challenges*, cit.

⁷¹ In particolare, nel momento in cui alcune amministrazioni locali hanno annunciato di voler consentire l'ingresso in alcuni spazi pubblici solo ai vaccinati, la reazione politica e mediatica è stata tale da indurre un rapido arretramento, con fonti giornalistiche che hanno documentato come il governo centrale abbia insistito sulla volontarietà della vaccinazione (Reuters, *After backlash, Beijing drops COVID vaccination mandate for crowded venues*, 8 luglio 2022).

⁷² Si veda quanto documentato da Human Rights Watch, *China's use of force and coercion to drive up its COVID-19 vaccination rate is not the answer*, 28 settembre 2021.

In parallelo, alcune Regioni hanno introdotto, dentro l'ecosistema degli "Health Codes", segnali visivi premiali per i vaccinati. La principale forma di incentivo è stata legata all'ottenimento degli indicatori specifici ("dorati") negli "Health Codes", necessari per una più agevole fruizione di servizi e per la circolazione⁷³. Questa forma di "gamification" non ha inciso sul colore del pass in senso stretto, ma ha funzionato come una forma di incentivo reputazionale, rendendo più agevole la fruizione e la circolazione in contesti ad alta densità di controlli⁷⁴.

Va tuttavia precisato che anche i non vaccinati potevano mantenere l'"Health Code" verde (e dunque circolare) se non intercettati da altri fattori di rischio (mobilità, esposizione, contatti, aree). In tale logica, la vaccinazione non fungeva da requisito unico e autosufficiente di accesso, ma come uno dei segnali – talora premianti – nell'ecosistema di classificazione del rischio⁷⁵.

Ciò ha rappresentato una differenza sostanziale rispetto all'esperienza italiana, soprattutto nella fase del "green pass" rafforzato, in cui l'accesso alle attività lavorative e sociali è stato riservato solo a vaccinati o guariti, visto che la vaccinazione o guarigione è diventata, per segmenti rilevanti della popolazione, una condizione direttamente abilitante in sé, non semplicemente un indicatore tra altri in un sistema di rischio⁷⁶.

L'opacità algoritmica costituisce forse la divergenza giuridicamente più rilevante rispetto al modello europeo. Mentre i criteri di validità del Green Pass erano trasparenti e verificabili, gli algoritmi che generavano i colori degli "Health Codes" sono rimasti completamente opachi: gli utenti non sapevano quali dati venivano utilizzati, come venivano ponderati, poiché il loro codice cambiava colore, né come contestare efficacemente una decisione algoritmica⁷⁷. Tale opacità strutturale ha reso possibile il "function creep"⁷⁸ documentato nel caso delle proteste bancarie nella provincia di Henan, nel giugno 2022, dove oltre mille cittadini hanno ricevuto, sul loro smartphone, il codice rosso non per ragioni sanitarie bensì per impedire loro di partecipare a manifestazioni contro il congelamento dei depositi bancari. Le successive sanzioni disciplinari ai funzionari coinvolti hanno confermato l'abuso sistemico ma non hanno rimosso la possibilità strut-

⁷³ Cfr. P. Della Posta, *La fine della politica dello zero-Covid in Cina: una valutazione*, in *ericaeconomia.it*, 14 gennaio 2023.

⁷⁴ Per un'analisi approfondita sul sistema del pass vaccinale cinese, cfr. H. Yu, *Living in the era of codes: a reflection on China's health code system*, in *Biosocieties*, 2022.

⁷⁵ Cfr. F. Yiyang, *Chinese Province Launches Gold 'Health Code' for Vaccine Recipients*, in *Sixth Tone*, 04 maggio 2021.

⁷⁶ Per una comparazione sui certificati vaccinali, cfr. B. Wang-Y. Ping, *A comparative analysis of Covid-19 vaccination certificates in 12 countries/regions around the world: rationalising health policies for international travel and domestic social activities during the pandemic*, in *Health Policy*, 126, 8, 2022, 755 ss.

⁷⁷ Sui diversi caratteri dei certificati vaccinali durante la pandemia, si veda l'interessante contributo di A. Alemanno-L. Bialasiewicz, *Certifying Health: The Unequal Legal Geographies of Covid-19 Certificates*, in *European Journal of Risk Regulation*, 2021, 2 ss.

⁷⁸ Cfr. A. Bernot-S. Trevaskes, *Containing Covid-19 through "surveillance creep" in China: Expansion of digital and physical surveillance*, in *Communicating Covid-19*, 2024. 325 ss.

turale di tale deriva⁷⁹.

Infine, la questione decisiva è la persistenza. La dismissione di strumenti nazionali di tracciamento della mobilità, come la “*Communication Itinerary Card*”, il 13 dicembre 2022, ha comportato lo spegnimento di tutti i canali (sms, web, *WeChat* e *Alipay*) e l’impegno degli operatori a cancellare i dati di percorso degli utenti, nel quadro dell’abbandono della politica “zero-Covid”. Tuttavia, l’esperienza ha poi documentato la tendenza, in varie città e province, a mantenere la struttura dell’“*Health Code*”, integrandola in servizi pubblici più ampi (sanità, trasporti, turismo, servizi amministrativi). Ad esempio, a Pechino, il “*Beijing Health Kit*” è confluito nel mini-programma “*JingTong*” (autenticazione e servizi sanitari e amministrativi), con un riuso dei dati di verifica delle identità sulla base del consenso⁸⁰. La persistenza di tale strumento ha configurato un processo di “*ratchet effect*”, teorizzato da R. Higgs⁸¹, in cui infrastrutture di sorveglianza temporaneamente giustificate dall’emergenza si consolidano in strumenti permanenti di governance amministrativa, potenzialmente riattivabili per future emergenze o estendibili ad altri scopi di controllo sociale⁸². Sotto questo profilo, il caso cinese illustra il rischio che misure eccezionali, una volta incorporate nelle architetture digitali e nelle prassi amministrative quotidiane, acquisiscano una propria inerzia istituzionale e resistano alla revoca formale, trasformandosi da strumenti emergenziali in dispositivi strutturali di governo della popolazione⁸³.

4. Il monitoraggio algoritmico della mobilità: tre paradigmi di *governance* dei dati in emergenza

In relazione al monitoraggio della mobilità, ciascun ordinamento qui considerato – UE, USA e Repubblica Popolare Cinese – ha costruito, giustificato e reso operativo il potere di tracciare algoritmicamente gli spostamenti individuali in nome della salute pubblica secondo logiche divergenti, che riflettono diverse concezioni della relazione tra dato, autorità pubblica e diritti. Nell’Unione europea è prevalsa una logica macro-epidemiologica basata su dati aggregati e soluzioni progettate per ridurre l’esposizione informativa dell’individuo; negli Stati Uniti ha pesato l’assenza di una cornice federale unitaria, con conseguente ricorso esteso a *dataset* commerciali e a *governance* contrattuale; in Cina, infine, il tracciamento è stato integrato

⁷⁹ Per i fatti di Henan, v. R. Boruta, *Henan protests: the “Chinese dream” in tatters?*, in *China Review*, 4, 2022.

⁸⁰ Per maggiori informazioni, v. [la pagina dedicata sul Foreign Affairs Office](#), in *wh.beijing.gov.cn*.

⁸¹ R. Higgs, *Crisis and Leviathan: Critical Episodes in the Growth of American Government*, Oxford, 1987.

⁸² Come osservato anche da G. Agamben, la creazione volontaria di uno stato di emergenza permanente è diventata una delle pratiche essenziali degli Stati contemporanei, dove il paradigma della sicurezza si è trasformato da eccezione a normale tecnica di governo. Così G. Agamben, *Lo Stato di eccezione*, Milano, 2003, 25 ss.

⁸³ Cfr. A. Bernot-S. Trevaskes, *Containing COVID-19 Through “Surveillance Creep” in China: expansion of Digital and Physical Surveillance*, cit., 321 ss.

negli strumenti amministrativi ordinari, con un altissimo grado di capillarità e una più limitata contestabilità delle decisioni automatizzate. È chiaro che la diversa combinazione di base legale, architetture tecniche e sedi di responsabilità ha inciso in modo diretto su trasparenza, fiducia sociale e *accountability*.

È altresì importante chiarire, sin d'ora, che dietro la definizione di “monitoraggio della mobilità” convivono – come ben chiarito nella Raccomandazione (UE) del 2020⁸⁴ – almeno due distinte tipologie di strumenti, che non devono essere sovrapposte: da un lato, l'uso di dati di mobilità (tipicamente aggregati) per monitorare flussi e valutare misure di contenimento; dall'altro, le tecnologie di *contact tracing*, implementate per notificare esposizioni senza ricostruire itinerari.

4.1. Dati di mobilità e prossimità “*privacy-preserving*” nel modello europeo

L'Unione europea ha affrontato il monitoraggio della mobilità adottando un paradigma di “*privacy-preserving epidemiology*”⁸⁵, caratterizzato da un costante bilanciamento tra efficacia epidemiologica e tutela dei diritti fondamentali. Tale approccio trova fondamento in una logica di trasparenza algoritmica istituzionalizzata, volta a rendere esplicite le condizioni sotto cui i dati vengono generati, trattati, e su un'*accountability* democratica multilivello, che colloca la legittimazione delle tecnologie di sorveglianza nel controllo pubblico e nella verificabilità *ex ante* ed *ex post* delle scelte tecniche, pur non mancando passaggi poco trasparenti, soprattutto, nella fase implementativa delle misure di sorveglianza a livello statale. In tale cornice, l'esame del caso europeo richiede la classificazione di due diverse linee di intervento, l'una relativa alla mobilità e l'altra, invece, alle misure di *contact tracing*.

Quanto all'impiego di dati di mobilità in forma aggregata e anonimizzata per finalità macro-epidemiologiche, la Raccomandazione (UE) n. 2020/518, dell'8 aprile 2020, ha rappresentato il punto di riferimento per misurare variazioni nei flussi, stimare l'efficacia delle restrizioni, supportare strategie di uscita dal *lockdown*. L'obiettivo è stato quello di delineare una strategia – pur non vincolante – diretta a orientare gli Stati membri verso un *toolbox* comune, basato sulla distinzione tra (i) applicazioni mobili e (ii) uso di dati di mobilità anonimizzati per finalità epidemiologiche, nel pieno rispetto dei principi del GDPR, tra cui necessità, proporzionalità e minimizzazione dei dati⁸⁶. Tale impostazione ha configurato un'*accountability* condivisa, imponendo agli Stati membri obblighi di motivazione pubblica rispetto alle scelte tecnologiche adottate. E la trasparenza non ha

⁸⁴ Raccomandazione (UE) 2020/518 della Commissione, 08 aprile 2020, relativa a un pacchetto di strumenti comuni dell'Unione per l'uso della tecnologia e dei dati al fine di contrastare la crisi Covid-19 e uscirne, in particolare per quanto riguarda le applicazioni mobili e l'uso di dati anonimizzati sulla mobilità.

⁸⁵ Definizione utilizzata da C. Troncoso-M. Payer-J.-P. Hubaux *et al.*, *Decentralized Privacy-Preserving Proximity Tracing*, in *arxiv.org*, 2020.

⁸⁶ Artt. 5, 6, 25, GDPR.

riguardato solo il codice sorgente, ma altresì le condizioni di produzione e trattamento del dato: chi genera il dataset, quali trasformazioni lo rendono anonimizzato, quali finalità epidemiologiche lo giustificano e quali limiti temporali ne impediscono la stabilizzazione oltre l'emergenza.

La seconda linea di intervento – diversa dal “monitoraggio della mobilità” in senso stretto – ha riguardato le tecnologie di *contact tracing* di prossimità. In tale contesto, le *Guidelines* 04/2020⁸⁷ dell'*European Data Protection Board* hanno rafforzato il nucleo “*privacy-preserving*” della risposta europea, raccomandando soluzioni basate su prossimità *Bluetooth* – atteso che la geolocalizzazione non è, in linea di principio, necessaria al tracciamento dei contatti – e richiedendo volontarietà, minimizzazione, limitazione della conservazione e valutazione d'impatto (DPIA).

Tale scelta “*privacy-preserving*” ha un significato ben preciso per quanto concerne la salvaguardia dei principi costituzionali dell'ordinamento europeo, in quanto ha reso il tracciamento compatibile con i diritti non riducendo a slogan il bilanciamento, ma disegnando architetture che impedissero *ex ante* la trasformazione di tali strumenti in sorveglianza dei percorsi individuali⁸⁸.

Tuttavia, la coerenza di questo approccio si è confrontata con una realtà tecnologica più complessa, che ha rivelato tensioni significative tra autonomia decisionale pubblica e potere infrastrutturale delle piattaforme private. Difatti, in parallelo all'azione pubblica, le *Exposure Notification APIs* di Apple e Google – annunciate il 10 aprile 2020 – non hanno solo risolto vincoli tecnologici strategici – in particolare per quanto concerneva la capacità di far funzionare lo scambio *Bluetooth* in background su iOS⁸⁹ – ma hanno altresì finito per delimitare il perimetro del “tecnicamente praticabile” per molte autorità sanitarie⁹⁰. In termini di *accountability* democratica, tale dinamica appena rappresentata descrive non tanto il potere esercitato da Apple e Google in relazione alla possibilità di accedere direttamente ai dati sanitari o di prossimità degli utenti – che l'architettura tecnologica ha impedito strutturalmente – quanto la capacità di definire i vincoli tecnici, gli standard di interoperabilità e le condizioni di possibilità entro cui le politiche pubbliche sono state chiamate ad operare nell'ecosistema degli *smartphone*⁹¹. L'esperienza britannica è istruttiva: la scelta iniziale per un impianto più centralizzato è stata abbandonata anche per limiti di performance legati al comportamento di iOS, con conseguente riposizionamento verso il modello Apple-Google⁹².

⁸⁷ Le “*Guidelines* 04/2020 on the use of location data and contact tracing tools in the context of the Covid-19 outbreak”, adottate il 21 aprile 2020.

⁸⁸ Una sintesi dell'architettura complessiva dell'UE può essere consultata in Commissione Europea, *Mobile applications to support contact tracing in the EU's fight against Covid-19: Common EU Toolbox*, 15 aprile 2020.

⁸⁹ Per ulteriori approfondimenti, v. V.L. Raposo, “*I'm right behind you*”: *Digital contact tracing under European Law*, in *Maastricht Journal of European and Comparative Law*, 29, 4, 2022.

⁹⁰ Parlamento europeo, *National COVID-19 contact tracing apps*, PE 652.711, 2020.

⁹¹ L. DeNardis, *Protocol Politics. The Globalization of Internet Governance*, The MIT Press, 2014.

⁹² G. Samuel-R. Sims, *The UK Covid-19 contact tracing app as both an emerging technology and public health intervention: The need to consider promissory discourses*, in *PubMed Central*, 2021.

Sul piano transfrontaliero, l'UE ha tentato di recuperare controllo pubblico attraverso l'interoperabilità dell'*European Federated Gateway Service*, attivato nell'ottobre 2020, per consentire alle app nazionali di “parlarsi” e gestire avvisi in mobilità intra-UE senza creare un archivio sanitario centralizzato. Tale elemento ha senz'altro rafforzato la dimensione “pubblica” dell'ecosistema, ma non ha eliminato la dipendenza infrastrutturale dall'OS mobile come livello abilitante.

Il nesso causale tra regole pubbliche europee preesistenti e *standard* tecnici a livello di sistema operativo forniti da attori privati ha effettivamente consentito un'inedita interoperabilità algoritmica transfrontaliera, rendendo tracciabile il funzionamento delle app e verificabili le scelte pubbliche, ma al prezzo di riconoscere ai *gatekeeper* un ruolo di co-regolatori nel definire il perimetro del tecnicamente possibile e del politicamente implementabile⁹³. Tale intricato connubio ha sollevato interrogativi strutturali sulla sovranità digitale europea⁹⁴ e sul controllo politico delle logiche computazionali che governano la sfera pubblica. La trasparenza, pur formalmente rispettata, si è così spostata dal piano della decisione alla dimensione tecnica dell'implementazione, ponendo la sfida – ancora aperta – di un'*accountability* democratica dell'interfaccia algoritmica⁹⁵.

4.1.1. *Contact Tracing* e sorveglianza della mobilità: le esperienze nazionali

Analogamente alle certificazioni digitali, anche per quanto concerne i sistemi di *contact tracing*, l'approccio europeo si è concretizzato in implementazioni nazionali che, pur condividendo il *framework* tecnico-normativo⁹⁶, hanno mostrato significative variazioni nell'adozione e nell'efficacia operativa.

La Germania ha offerto l'esempio più coerente con l'ideale europeo “*privacy-preserving*”: utilizzo di Bluetooth (non GPS), logiche di prossimità, caricamento volontario delle chiavi in caso di positività e ampia documentazione tecnica pubblica. La “*Corona-Warn-App*”, basata sulla piattaforma frutto della collaborazione tra Google e Apple, sviluppata da SAP e *Deutsche Telekom* su commissione del Governo federale, è stata lanciata il 16 giugno 2020⁹⁷. L'applicazione ha raggiunto decine di milioni di *download* cumulativi entro la fine del 2023, ma – come accade tipicamente in questi

⁹³ In merito al ruolo di co-regolatori dei *gatekeeper*, v. G. Kerber, *Updating Competition Policy for the Digital Economy?*, in *Antitrust Bulletin*, 2019, 64, 3 ss.

⁹⁴ In tema, v. A. Alpini, *La sovranità digitale europea*, in *Rivista di Scienze della Comunicazione e di Argomentazione Giuridica*, XIV, 2, 2022, 144 ss.; M. Santaniello, *Sovranità digitale e diritti fondamentali: un modello europeo di Internet governance*, in *Rivista Italiana di Informatica e Diritto*, 1, 2022, 48 ss.

⁹⁵ Si veda l'interessante documento del Parlamento europeo, *Digital sovereignty for Europe*, 2020.

⁹⁶ Decisione (UE) 2020/1023, 15 luglio 2020.

⁹⁷ Per ulteriori approfondimenti sul processo di implementazione del “*Corona-Warn-App*”, v. M.F. Enaya-T. Klingbeil-J. Krüger *et al.*, *A case study on the development of the German Corona-Warn-App*, in *Journal of Systems and Software*, 213, 2024.

strumenti – con un tasso di utenti attivi significativamente inferiore⁹⁸.

Tuttavia, accanto a una documentata ed evidente trasparenza tecnica, l'*accountability* sostanziale ha risentito di non poche limitazioni, in quanto fortemente condizionato dalla capacità del sistema sanitario di trasformare una notifica in azione e, dunque, da una catena che non è riducibile al solo codice⁹⁹.

L'Italia ha adottato un approccio analogo con "Immuni", sviluppata da "Bending Spoons" e resa disponibile dal 1° giugno 2020¹⁰⁰. Il tratto distintivo, sul piano costituzionale, è stato l'innesto precoce di un'*accountability* istituzionale *ex ante*, a seguito dell'intervento con cui l'Autorità Garante ha condizionato l'impianto con prescrizioni funzionali alla limitazione dei dati e alla prevedibilità del trattamento, rafforzando la trasparenza procedurale¹⁰¹.

Dal 19 ottobre 2020, Immuni è stata integrata nel Gateway europeo EFGS (*European Federation Gateway Service*) permettendo l'interoperabilità transfrontaliera delle notifiche di esposizione¹⁰². L'inclusione in questa infrastruttura ha implicato la trasparenza delle specifiche tecniche relative allo scambio di dati tra Stati, rafforzando l'*accountability* multilivello.

Nonostante le solide premesse progettuali e normative, l'efficacia operativa di "Immuni" ha risentito di alcune criticità che hanno inciso sia sull'efficienza del sistema sia sulla responsabilizzazione degli attori coinvolti¹⁰³. Tra i principali fattori limitanti si segnalano la bassa penetrazione nella popolazione, soprattutto nelle fasi iniziali, le difficoltà tecniche su dispositivi iOS che impedivano il funzionamento dell'app in background, nonché la scarsa integrazione con i sistemi sanitari regionali, manifestatasi in ritardi nella comunicazione dei codici di sblocco agli utenti positivi¹⁰⁴. Tali criticità hanno messo in luce una parziale frammentazione della responsabilità operativa tra sviluppatori, Ministero della Salute, comitati tecnici e autorità regionali, relativizzando l'efficacia dell'*accountability* sostanziale.

Operativamente, in Italia, i *download* di "Immuni" hanno superato i 10 milioni già a dicembre 2020, arrivando a circa 22 milioni complessivi al momento della dismissione (27 dicembre 2022). Al 31 dicembre 2021,

⁹⁸ R. Müller-Török-A. Prosser, *The Corona Warning App of the German Federal Government. How perceived data protection issues hindered any effectiveness*, in *Smart Cities and Regional Development Journal*, 2021, 23 ss.

⁹⁹ In tema di accettabilità del dispositivo da parte dei cittadini, v. F. Beierle-U. Dhakal-C. Cohrdes *et al.*, *Public Perception of the German COVID-19 Contact-Tracing App Corona-Warn-App*, 2021, in *arXiv.org*.

¹⁰⁰ Per un'analisi dettagliata delle fonti legali, delle caratteristiche e limiti dell'app "Immuni", v. M. Plutino, "Immuni". *Un'exposure notification app alla prova del bilanciamento tra tutela dei diritti e degli interessi pubblici*, in questa Rivista, 2, 2020, 173 ss.

¹⁰¹ Garante per la protezione dei dati personali, *Valutazione d'impatto sulla protezione dei dati personali presentata dal Ministero della Salute relativa ai trattamenti effettuati nell'ambito del sistema di allerta Covid-19 denominato "immuni". Nota sugli aspetti tecnologici*, 3 giugno 2020.

¹⁰² Decisione (UE) 2020/1023.

¹⁰³ Per una lettura approfondita dei meccanismi di funzionamento, v. R. Berti-A. Longo-S. Zanetti, *Immuni, cos'è e come funziona l'app italiana coronavirus*, in *agendadigitale.eu*.

¹⁰⁴ E. Santoro, *App anti-covid, le ragioni del flop: ecco cosa si poteva fare meglio (non solo in Italia)*, in *agendadigitale.eu*, 1 dicembre 2020; Id., *L'app Immuni è ora disponibile: tanti ancora i dubbi sulla sua efficacia*, 1 dicembre 2020.

le notifiche risultavano essere 143.956 e i casi identificati 44.880¹⁰⁵. Pur rappresentando dati di trasparenza statistica, non sono mai stati pubblicati studi indipendenti riguardo all'impatto epidemiologico reale dell'app, né sono stati messi a disposizione dati dettagliati o modelli predittivi per una verifica esterna. Inoltre, sono mancate procedure formali per contestare le notifiche ricevute o per richiedere una revisione di eventuali errori algoritmici, rappresentando così una limitazione nella piena realizzazione di un'*accountability* partecipata e sostanziale.

Diversamente da Germania e Italia, la Francia ha inizialmente tentato un approccio centralizzato con "*StopCOVID*", rifiutando le GAEN di Apple/Google e sviluppando un'architettura proprietaria basata sul protocollo "*ROBERT*" (*ROBust and privacy-presERving proximity Tracing*), sviluppato dall'*Institut National de Recherche en Informatique et en Automatique*. Tale scelta, motivata da considerazioni di sovranità digitale, ha però creato problemi tecnici significativi, in quanto, su dispositivi iOS, l'app non funzionava in *background*, condizionando anche l'adozione che è rimasta molto bassa (inferiore a 3 milioni di utenti)¹⁰⁶.

Di fronte al fallimento, nell'ottobre 2020, il Governo francese ha effettuato una svolta strategica, rilanciando l'app come "*TousAntiCovid*", ma senza migrare a GAEN: l'architettura è rimasta centralizzata, rendendo tale app l'eccezione europea rispetto al modello decentralizzato prevalente¹⁰⁷.

Dal 2021 l'app è stata progressivamente ri-funzionalizzata, acquisendo al suo interno anche il "*pass sanitaire*", con un forte aumento dei *download*: oltre 39,4 milioni di registrazione al 1° gennaio 2022 secondo il *Rapport d'activité* del Ministero della Salute. Tuttavia, meno del 7% degli utenti ha ricevuto poi effettivamente una notifica di contatto, segno evidente che l'app è stata usata soprattutto come strumento per mostrare il "*pass sanitaire*", mentre la parte di *contact tracing* è rimasta minoritaria¹⁰⁸.

Sul versante della trasparenza, Francia ha pubblicato codice e documentazione tecnica del protocollo e la CNIL ha emesso pareri ricorrenti sull'app. Ciò ha reso "visibile" l'architettura e parte delle logiche, ma il modello centralizzato ha mantenuto sul server pubblico il controllo del calcolo e della gestione dei contatti e delle politiche del rischio, accentuando l'asimmetria informativa rispetto alle soluzioni "on-device" (GAEN).

Anche sul versante dell'*accountability*, la centralizzazione ha reso più chiaro la responsabilità dei soggetti istituzionali, ma ha anche posto interrogativi sulla proporzionalità del trattamento rispetto alla finalità di *contact tracing* e, soprattutto, sull'effetto reale dell'app nel contenimento dei contagi.

In estrema sintesi, Germania e Italia mostrano un paradigma "europeo" relativamente lineare, mentre la Francia evidenzia che l'architettura può

¹⁰⁵ [Repository GitHub](#), in [github.com](#).

¹⁰⁶ CNIL, *Délibération n° 2020-046 du 24 avril 2020 portant avis sur un projet d'application mobile dénommée «StopCovid»*, Avis n° 20006919.

¹⁰⁷ E. Schultz-R. Touzani-J. Macini *et al.*, *From contact tracing to Covid-19 pass holder; the tortured journey of the French TousAntiCovid contact tracing app*, in HAL, open science, 2024; A. Blasimme-A. Ferretti-E. Vayena, *Digital Contact Tracing Against Covid-19 in Europe: Current Features and Ongoing Developments*, in *Digital Health*, 2021.

¹⁰⁸ Ministère des Solidarités et de la Santé, *Rapport d'activité, TousAntiCovid du 2 juin 2020 au 30 novembre 2021*.

essere resa più “statale” senza che ciò produca tuttavia un guadagno netto in legittimazione, perché l’adozione e l’efficacia restano politicamente e socialmente mediate.

4.1.2. Dati aggregati di mobilità e trasparenza algoritmica: alcuni casi critici nei Paesi dell’UE

Parallelamente all’esperienza dei sistemi di *contact tracing*, molti Stati membri dell’UE hanno utilizzato dati aggregati di mobilità (fonti telco, statistiche pubbliche, dataset) al fine di alimentare modelli previsionali a supporto della valutazione delle restrizioni territoriali (“zone colorate”)¹⁰⁹. In questa famiglia di strumenti, tuttavia, la trasparenza è apparsa più fragile, visto che l’intero processo decisionale è stato allocato in un’infrastruttura analitica che ha operato “a monte” della decisione.

Nel caso italiano, il legislatore ha autorizzato l’uso di dati in forma anonima e aggregata in chiave emergenziale e di innovazione digitale. Tali dati, pur formalmente processati in conformità alle linee guida EDPB, hanno contribuito ad alimentare – unitamente a un *set* di indicatori epidemiologici e di capacità sanitaria legalmente prefissati¹¹⁰ – modelli previsionali che hanno condizionato il godimento di diritti fondamentali come la libertà di circolazione, lavoro ed attività economica. La questione ha assunto rilievo critico in considerazione del fatto che, a differenza delle app di *contact tracing* sottoposte a DPIA pubblicate e *audit* indipendenti, per l’utilizzo dei dati degli operatori telefonici non è stata condotta alcuna DPIA che valutasse rischi di re-identificazione, proporzionalità dell’intervento e misure di mitigazione, determinando un’opacità procedurale e un *deficit* di trasparenza algoritmica.

Inoltre, la mancata pubblicazione dei modelli previsionali e l’assenza di revisione scientifica indipendente hanno comportato una grave carenza di *accountability* epistemica, ossia l’impossibilità per la cittadinanza e la comunità scientifica di verificare la solidità e l’affidabilità delle previsioni che hanno inciso pesantemente sul godimento di importanti diritti fondamentali. Sul piano della responsabilità pubblica, la decisione in capo a numerosi soggetti (operatori telefonici, Ministero della Salute, comitati tecnici, Regioni) ha poi determinato una diluizione della responsabilità, complicando l’identificazione di un soggetto unico imputabile per eventuali errori, con assenza di meccanismi di contestazione o revisione delle decisioni algoritmiche.

Anche la Spagna ha fatto ricorso all’utilizzo dei dati aggregati e anonimizzati di mobilità, dimostrando un approccio parzialmente più trasparente rispetto all’Italia. L’*Instituto Nacional de Estadística* (INE) ha utilizzato dati aggregati degli operatori mobili per produrre l’*Estudio de Movilidad a partir*

¹⁰⁹ M. Vespe-U. Minora-S. Iacus *et al.*, *Mobility and Economic Impact of Covid-19 Restrictions in Italy using Mobile Network Operator Data*, Publications Office of the European Union, 26 maggio 2021.

¹¹⁰ D.M. 30-4-2020; D.P.C.M. 3 novembre 2020.

de la telefonía móvil¹¹¹, pubblicando report settimanali sui *pattern* di spostamento della popolazione durante le diverse fasi di *lockdown*. Sono state inoltre rese disponibili *dashboard* con visualizzazioni aggregate per province e documentazione tecnica sulla metodologia di aggregazione.

Dopo la fine dello *estado de alarma*, nel maggio 2021, il *Ministerio Transportes, Movilidad y Agenda Urbana* (MITMA) ha avviato, a partire dal 1° gennaio 2022, uno studio continuo basato su *Big Data* e registri anonimizzati degli operatori e registri anonimizzati, con pagine metodologiche ufficiali¹¹². È evidente che la persistenza dell'infrastruttura come paradigma ordinario dal 2022 (oltre l'emergenza) segnala un riuso strutturale dei flussi di mobilità; ciò richiede basi giuridiche chiaramente esplicitate, valutazioni di proporzionalità aggiornate e governance specifica e dedicata, al fine di evitare fenomeni di “*function creep*”¹¹³, dove dati raccolti per emergenza sanitaria migrano verso usi amministrativi ordinari senza adeguato scrutinio democratico¹¹⁴.

In Germania, il “*COVID-19 Mobility Project*” – promosso dal Robert Koch Institut in collaborazione con il gruppo Brockmann della Humboldt-Universität – ha monitorato i flussi di mobilità tramite dati aggregati e anonimizzati derivati dalla telefonia mobile (Deutsche Telekom e Telefónica)¹¹⁵, con dati analizzati e aggregati da Teralytics¹¹⁶, al fine di monitorare l'efficacia delle restrizioni e informare le decisioni di politica sanitaria a livello dei *Länder*. La documentazione pubblicata del progetto evidenzia che i *dataset* consistono in conteggi di spostamenti intra e inter-circoscrizionali, costruiti attraverso soglie minime e aggregazione spazio-temporale, in modo da escludere la riconducibilità a persone fisiche¹¹⁷.

¹¹¹ Instituto Nacional de Estadística, *Estudios de movilidad a partir de la telefonía móvil*.

¹¹² Ministerio de Transportes y Movilidad Sostenible, *Estudio de la movilidad con Big Data*.

¹¹³ Il fenomeno del “*function creep*” è ampiamente discusso in dottrina come un cambiamento graduale e spesso impercettibile nell'uso dei sistemi di trattamento dati che, originariamente concepiti per scopi specifici, vengono impiegati per finalità ulteriori e non sempre legittimate. Koops definisce il “*function creep*” come una trasformazione che, pur sembrando innovazione, comporta rischi significativi per i diritti fondamentali e la trasparenza, richiedendo un dibattito pubblico e misure di contenimento efficaci. In tale prospettiva, molti sottolineano che il principio di limitazione delle finalità del GDPR può svolgere un ruolo cruciale nel contrastare tale fenomeno, auspicando valutazioni continue e trasparenti sull'utilizzo dei dati raccolti per specifici scopi. V. B.-J. Koops, *The Concept of Function Creep*, in *Law, innovation and technology*, 13, 2020; S. Fantin-P. Vogiatzoglou, *Purpose Limitation by Design as a Counter to Function Creep and System Insecurity in Police Artificial, United Nations Interregional Crime and Justice Research Institute (UNICRI), UNICRI Special Collection on AI*, 2020; M. Naarttijärvi, *Function creep, altered affordances, and safeguards rollbacks*, in *verfassungsblog.de*, 11 aprile 2022.

¹¹⁴ M. Madianou, *A Second-Order Disaster? Digital Technologies During the Covid-19 Pandemic*, in *Social Media+Society*, 2020.

¹¹⁵ Ulteriori informazioni su metodi di anonimizzazione, fonti dei dati e strumenti di analisi, sono reperibili all'url covid-19-mobility.org.

¹¹⁶ Informazioni più dettagliate sulla capacità dell'azienda di fornire dati di mobilità e modelli predittivi basati su oltre 30 milioni di *device* in Germania, sono reperibili all'url teralytics.net/matrix/mobility-trends.

¹¹⁷ L'Ufficio Federale di Statistica tedesco (*Statistisches Bundesamt*) pubblicava dati aggregati di mobilità basati su telefonia mobile, sottolineando le procedure di anonimizzazione sviluppare in collaborazione con il Garante della *privacy* (BDI). La reportistica, tuttavia, ha

Cionondimeno, dal punto di vista della trasparenza algoritmica, il processo di trattamento dei dati ha incluso componenti proprietarie che non sono state integralmente documentate al pubblico, limitando la verificabilità delle metriche rispetto a progetti integralmente “open”. Al tempo stesso, il portale del progetto ha fornito descrizioni metodologiche, definizioni operative e riferimenti specifici, garantendo una trasparenza di livello macro sui risultati esposti. In breve, si è avuta una buona trasparenza descrittiva sul versante istituzionale, a fronte di una opacità residua nelle fasi proprietarie di pre-trattamento¹¹⁸.

Sul piano dell'*accountability*, la presenza di attori privati lungo tutto il processo di trattamento dei dati ha segmentato le responsabilità. Ne è derivata l'esigenza di una allocazione chiara dei ruoli e, idealmente, di *audit* indipendenti su tutte le varie fasi di monitoraggio. Diversamente dalla “*Corona-Warn-App*” – sottoposta, come già visto, a verifiche pubbliche coordinate dal BSI – per i flussi di dati relativi alla mobilità non risultano *audit* pubblicati di portata comparabile, con conseguente asimmetria di garanzie tra strumenti parimenti impiegati a fini di sanità pubblica.

La competenza in materia sanitaria dei *Länder* ha prodotto pratiche eterogenee. In alcuni contesti, gli enti regionali hanno stipulato accordi propri per analisi settoriali – ad esempio, in Baviera – tramite *Bayerisches Zentrum für Tourismus*, con dati di Telefónica e Teralytics – altrove si è privilegiata una documentazione metodologica e un'offerta di open data più strutturata, come nel caso del Baden-Württemberg¹¹⁹, non sempre però specifiche rispetto al tracciato Covid-19. Il risultato è stata una trasparenza variabile tra i vari *Länder*, che ha reso più complesso uno scrutinio pubblico uniforme. L'insieme di tali esperienze mette in evidenza un divario strutturale tra *compliance* formale e *accountability* sostanziale. Se i modelli e le catene decisionali restano non replicabili e se i livelli proprietari non sono sottoposti a verifiche esterne, l'uso di dati aggregati rischia di produrre una “zona grigia” di governo algoritmico difficilmente contestabile.

4.2. Il modello statunitense tra frammentazione federale e *surveillance laundering*

L'esperienza statunitense di gestione della mobilità digitale in pandemia ha incarnato – similmente a quanto constatato nella gestione delle certificazioni digitali – un modello privo di coordinamento federale e fortemente dipendente da soluzioni frammentate nate da *partnership* pubblico-private¹²⁰. In un quadro privo di una cornice federale organica di tutela dei dati

evidenziato limiti metodologici, mancando di pubblicare dettagli sugli algoritmi specifici. Cfr. maggiori dettagli all'url [Mobility indicators based on mobile network data](#), in [destatis.de](#).

¹¹⁸ Maggiori informazioni sono reperibili sul sito Internet dell'iniziativa in [covid-19-mobility.org](#).

¹¹⁹ Per maggior dettagli cfr. [Mobilitätsverhalten während der Corona-Pandemie](#) in [vm.baden-wuerttemberg.de](#)

¹²⁰ Cfr. l'attenta ricostruzione di L. Breeher-A. Boon-C. Hainy *et al.*, [A Framework for Sustainable Contact Tracing and Exposure Investigation for Large Health Systems](#), in *Mayo Clin Proc.*, 2020.

personali – il *Privacy Act* del 1974 copre ambiti federali limitati e l'*Health Insurance Portability and Accountability Act* (HIPAA) non si estende ai *data broker* – la tutela costituzionale passa per una combinazione di fattori che includono, tra l'altro, Quarto Emendamento, “*third-party doctrine*” (come rimodellata da “*Carpenter v. United States*”¹²¹) e strumenti *ex post* di controllo (FOIA), più che per garanzie *ex ante* di tipo “*privacy by design*”.

Durante la pandemia, le agenzie sanitarie a livello federale, statale e locale hanno fatto ampio affidamento su dati forniti da società private, inizialmente a titolo gratuito e successivamente acquisiti mediante contratti pubblici. Tale prassi ha generato un regime ibrido “pubblico-privato” di sorveglianza, caratterizzato da elevata potenza analitica ma, al contempo, da significative problematiche di *accountability* democratica.

Un primo canale è stato rappresentato dai “*Google Covid-19 Community Mobility Reports*” e dagli “*Apple Mobility Trends Reports*”, utilizzati in maniera estesa nella ricerca epidemiologica per misurare variazioni negli spostamenti e nei tempi di permanenza in varie zone del Paese¹²². Tali strumenti, pur offrendo un'elevata utilità, spostano una porzione significativa di strumenti di controllo della vita dei cittadini su un terreno privato: metriche, baseline, policy di accesso sono definiti unilateralmente dalle piattaforme, sulla base di standard tecnici proprietari.

Un secondo canale – più sensibile sul piano costituzionale – riguarda i dati commerciali di localizzazione provenienti dall'ecosistema pubblicitario “*Software Development Kit*” (SDK), incorporati in migliaia di *app consumer*, *data broker*, che le autorità non “raccolgono” direttamente ma acquisiscono mediante contratti e appalti. Essi hanno offerto una granularità spazio-temporale superiore ai dati aggregati europei, abilitando analisi sofisticate sull'efficacia delle misure restrittive fino a livello di singoli quartieri o di singoli punti di interesse. Ma, come documentato da indagini giornalistiche e rapporti di organizzazioni per i diritti civili¹²³, la *governance* di questi flussi informativi si è retta soprattutto su clausole contrattuali e termini di servizio, sottraendo scelte sostanziali ai processi democratici, con ricadute significative in termini di rispetto dell'*accountability*. Basti citare, su scala federale, la costruzione di grandi piattaforme di integrazione dati come “*HHS Protect*”, descritto come un *data warehouse* in grado di combinare centinaia di dataset pubblici e privati a supporto delle decisioni emergenziali e associato in più ricostruzioni a contratti e tecnologia *Palantir* per la gestione e l'analisi dei dati sanitari e logistici¹²⁴. È chiaro che allorché l'infrastruttura di data governance è costruita prevalentemente per via contrattuale e tecnicamente mediata da fornitori privati, l'*accountability* tende a frammentarsi tra fasi di *procurement*, regole sulla qualità del dato e modelli analitici, rendendo più difficile attribuire responsabilità politiche e

¹²¹ US Supreme Court, *Carpenter v. United States*, 585 U.S. (2018).

¹²² K. Klise-W. Beyeler-P. Finley *et al.*, *Analysis of mobility data to build contact network for Covid-19*, in *PLoS One*, 16, 4, 2021.

¹²³ Electronic Frontier Foundation, *Digital Rights During the Covid-19 Pandemic, Special Report Series*, 2020-2021; ACLU, *Principles for Technology-Assisted Contact-Tracing, White Paper, White Paper*, aprile 2020.

¹²⁴ Center for Public Integrity, *New, secretive data system shaping federal pandemic response*, 2022.

giuridiche per errori, distorsioni o fenomeni di “*mission creep*”.

La dipendenza da fornitori commerciali ha altresì generato asimmetrie territoriali nell’accesso a capacità analitiche avanzate, favorendo gruppi con maggiori risorse finanziarie e relazioni consolidate con il settore tecnologico, mentre l’assenza di una legge federale generale sulla *privacy* ha spostato il controllo sui contratti e sul controllo *ex post* del *Freedom of Information Act*, con frequente contenzioso sull’estensione legittima su finalità e ampiezza d’uso, che vanno dal monitoraggio del rispetto dei coprifuochi al tracciamento degli accessi nelle scuole e presso i centri vaccinali¹²⁵.

Particolarmente significativo è stato il caso dell’accesso, da parte dei *Centers for Disease Control and Prevention*, ai dati aggregati di localizzazione forniti dal *data broker* “*SafeGraph*”, acquisiti per analizzare *pattern* di mobilità su larga scala¹²⁶. I documenti resi pubblici hanno rilevato che l’uso prospettato non era limitato al solo contenimento del contagio, ma includeva anche applicazioni più ampie, rendendo visibile il rischio di slittamento funzionale¹²⁷.

La vicenda ha sollevato questioni di legittimità “di sistema”, alimentando il dibattito sul “*surveillance laundering*”, ossia l’acquisizione indiretta, da parte di autorità pubbliche, di capacità di sorveglianza controverse, mediante l’intermediazione di soggetti privati che operano sotto regimi di *privacy* più permissivi¹²⁸.

Nel contesto pandemico, il controllo pubblico su tali strumenti ha ope-

¹²⁵ Si veda l’interessante e approfondito studio della Johns Hopkins Bloomberg School of Public Health, *A National Plan to Enable Comprehensive Covid-19 Case Finding and Contact Tracing in the US*, 2020.

¹²⁶ J. Cox, *CDC Tracked Millions of Phones to see if Americans Followed Covid Lockdown Orders*, 3 maggio 2022.

¹²⁷ Nel pieno della prima ondata pandemica, nell’aprile 2020, il Governo del Distretto di Columbia ha accettato un’offerta “pro bono” dal *data broker* Veraset per accedere – in via sperimentale – ai propri flussi di dati di localizzazione a livello di dispositivo relativi all’area metropolitana di Washington. I *files* relativi ai dati di tracciamento venivano consegnati con un ritardo compreso tra 24 e 72 ore rispetto all’evento, consentendo un’analisi quasi in tempo reale sui *pattern* di spostamento. Pochi giorni dopo, il 9 aprile 2020, le parti hanno stipulato un “*Data Access Agreement*” non esclusivo per “scopi di ricerca”, che ribadiva la titolarità dei dati in capo a Veraset, prevedendo limiti rigorosi di riuso e di trasferimento a terzi, nonché un obbligo di cancellazione a fine periodo. Nel complesso, il caso in parola esemplifica una traiettoria tipica della fase emergenziale negli Stati Uniti, con un avvio di “*trial*” di dati commerciali, successiva integrazione nelle infrastrutture analitiche pubbliche e, solo in un secondo momento, valutazione di efficacia dell’operazione. Al contempo, il caso illumina due snodi di principio. Da un lato, la capacità dei *dataset* di mobilità di fornire una granularità decisionale inedita (fino a livello di singolo quartiere) e, dall’altro, i limiti strutturali alla trasparenza esterna derivanti sia dalla classificazione interna sia dall’architettura contrattuale. Secondo ricostruzioni giornalistiche, il progetto avrebbe incluso miliardi di record, con aggiornamenti frequenti e copertura estesa all’intera area metropolitana di D.C., Virginia e Maryland, confermando il ruolo dell’esperimento come banco di prova delle relazioni pubblico-private nella *governance* digitale dell’emergenza. Per un’attenta ricostruzione del caso, v. B. Cyphers, *Data Broker Veraset Gave Bulk Device-Level GPS Data to DC Government*, in *Electronic Frontier Foundation*, 10 novembre 2021; D. Harwell, *Data broker shared billions of location records with District during pandemic*, in *Washington Post.com*, 10 novembre 2021.

¹²⁸ N. Richards-W. Hartzog, *A Duty of Loyalty for Privacy Law*, in *Washington University Law Review*, 97, 2019, 1663 ss.

rato in larga parte *ex post*. Difatti, molte informazioni su contratti, *dataset* utilizzati e finalità effettive sono emerse tramite richieste FOIA, audizioni congressuali e inchieste giornalistiche, più che attraverso meccanismi *ex ante* di valutazione d'impatto e di legittimazione democratica. Ciò ha reso la trasparenza non tanto un prerequisito di validità del trattamento, quanto un rimedio informativo tardivo, disperso lungo una catena di procedimenti e di attori.

In chiave comparata, il modello “USA” non appare semplicemente meno protettivo, ma diverso nel suo baricentro costituzionale. Laddove l'UE tende a legittimare l'impiego dei dati attraverso vincoli generali, valutazioni d'impatto e *privacy by design*, negli Stati Uniti, una quota rilevante della disciplina si sposta su piani meno pubblici e più contingenti: contratti di fornitura, standard industriali, meccanismi di controllo *ex post*, sullo sfondo di una giurisprudenza del Quarto Emendamento ancora impegnata a ridefinire la protezione dei dati di localizzazione nell'economia dei *data broker*. Il risultato, nel monitoraggio epidemiologico della mobilità, è stato ambivalente. Da un lato, un'elevata capacità di misurare e modellizzare i comportamenti ma, dall'altro, una metamorfosi più opaca di questa sorveglianza in decisioni pubbliche pienamente spiegabili e contestabili.

4.2.1. L'esperienza di *contact tracing* negli Stati Uniti tra *standard* privati e frammentazione dell'*accountability*

Diversamente dal quadro europeo, gli Stati Uniti non hanno sviluppato un'unica app federale di *contact tracing* né un'infrastruttura pubblica comparabile, per omogeneità e “presa istituzionale, ai dispositivi europei. La conseguenza è stata la diffusione, a livello statale, di strumenti di varia natura quanto a funzionamento e tracciabilità, spesso con un ruolo decisivo di attori privati nel *design* e nell'operatività quotidiana.

Il sistema di tracciamento è stato avviato a seguito della proposta congiunta Apple/Google di una tecnologia di “*exposure notification*”, basata su *Bluetooth*, che le autorità sanitarie hanno potuto incorporare nelle proprie applicazioni. Tale scelta ha poi legittimato l'esito tipicamente statunitense, ossia decentramento politico-istituzionale (ogni Stato ha deciso se e come adottare tale strumento) combinato con una standardizzazione tecnologica privata (l'ecosistema iOS/Android). Da ciò è derivato un modello nel quale il perimetro del “possibile” è, in parte, predefinito da vincoli tecnici e parametri di interoperabilità dettati dalle piattaforme private, mentre la legittimazione e l'uso concreto sono rimasti frammentati in ordinanze sanitarie, scelte amministrative e prassi settoriali¹²⁹.

Sul piano dell'adozione, le prassi nazionali hanno evidenziato un processo scaglionato, con 26 Stati che hanno implementato l'app di “*exposure notification*” in circa dieci mesi, a partire dall'agosto 2020 con lo Stato della

¹²⁹ E. Jang-J. Baik-K. Fischer, *Contact-Tracing as Boundary of Pandemic Governance: The State-by-State to Contain the Spread of COVID-19 in the United States*, in *International Journal of Communication*, 17, 2023, 1737 ss.

Virginia, senza però che si consolidasse un programma nazionale unitario¹³⁰. Allo stesso tempo, al fine di evitare che ogni singolo Stato dovesse costruire un proprio “*key server*”, nonché per favorire l’interoperabilità, un ruolo centrale è stato assegnato all’*Association of Public Health Laboratories*” (APHL), con la creazione di un “*National Key Server*”, che ha favorito una circolazione più fluida delle notifiche in mobilità interstatale. Su tale punto emerge un tratto che la letteratura ha messo a fuoco in maniera chiara, ossia che le app di *contact tracing* americane hanno funzionato come “*boundary objects*” della *governance* pandemica, ossia oggetti capaci di mettere in contatto istituzioni, imprese e utenti, ma anche di spostare la regolazione dal diritto pubblico a standard tecnici, *policy* di piattaforma e contratti¹³¹.

Ed è proprio in tale scenario che trasparenza e *accountability* si sono scomposti e frammentati. Se la trasparenza è risultata elevata sul piano tecnologico e infrastrutturale, essa si è indebolita proprio dove l’esperienza europea ha tentato (pur imperfettamente) di costruire controlli *ex ante*, ossia sui criteri di *policy*, soglie, esenzioni, condizioni d’uso, che negli USA sono risultati dispersi tra ordinanze statali e locali, linee guida, protocolli di settore e regole operative dei *provider*.

Sul piano dei rimedi l’asimmetria è stata ancora più netta. L’utente che ha contestato un esito si è spesso perso all’interno di una filiera caratterizzata da molteplici centri decisionali (autorità sanitaria, provider, app, datore di lavoro), che hanno finito con l’elevare i costi di contestazione e ridurre l’effettività della tutela, in un momento in cui la tempestività non ha rappresentato solo un dettaglio procedurale, ma parte integrante del diritto alla protezione sanitaria¹³².

In sintesi, l’esperienza statunitense di *contact tracing* mostra come un’architettura *privacy-preserving* non garantisce di per sé un’*accountability* solida, non tanto perché manchino i principi di riferimento, quanto perché la responsabilità operativa e regolativa si è dispersa in regole prevalentemente procedurali o contrattuali, difficilmente sottoponibili a controlli *ex ante* comparabili a quelli che, almeno tendenzialmente, caratterizzano il paradigma europeo.

4.3. Il modello cinese tra sorveglianza sanitaria e controllo sociale

Il modello cinese di monitoraggio della mobilità digitale durante la pandemia si è distinto per l’integrazione di strumenti digitali dentro la *governance* amministrativa ordinaria, secondo uno schema logico in cui i dati di spostamento, l’accesso ai luoghi e la classificazione del rischio convergevano in un’unica architettura operativa.

Diversamente dall’esperienza europea – in cui la raccolta dei dati di mobilità è rimasta tendenzialmente ancorata a finalità macro-epidemiologiche

¹³⁰ *Ibid.*

¹³¹ *Ibid.*

¹³² Per ulteriori approfondimenti, v. B. Levy-M. Stewart, *The Evolving Ecosystem of Covid-19 Contact tracing Applications*, in *Harvard Data Science Review*, 3, 3, 2021.

e a vincoli *ex ante* – la risposta cinese ha combinato, da un lato, la perimetrazione territoriale delle aree a rischio e, dall’altro, l’introduzione di filtri individuali di accesso costruiti su infrastrutture digitali pervasive¹³³.

Sul primo versante, la gestione della mobilità è passata, in particolare, attraverso strumenti di “*risk zoning*”, aventi effetti amministrativi immediati, in termini di restrizioni alla circolazione e controllo selettivo degli spazi. I protocolli di prevenzione e controllo emanati dalla *China CDC* contemplavano, soprattutto per le aree classificate ad alto rischio, restrizioni alla mobilità locali e interregionali¹³⁴. Nella prassi, tale prassi si è tradotta in un’amministrazione “a grana fine” (*fine-grained governance*) delle comunità territoriali, attraverso la distinzione tra “*sealed zone*”, “*controlled zone*” e “*precautionary zone*”, che attribuiva alle autorità locali il potere di modulare in tempo reale la libertà di movimento, con implementazione resa effettiva dalle piattaforme digitali.

Sul secondo versante, le misure di *contact tracing* sono state assicurate attraverso la convergenza tra “*Healths Codes*” locali – inseriti nelle piattaforme di *WeChat* e *Alipay* – e “*Communication Itinerary Card*”. In particolare, quest’ultima – alimentata da dati forniti dagli operatori di telecomunicazioni – ha funzionato come filtro interregionale basato sulla cronologia delle zone visitate, diventando requisito implicito per l’accesso a servizi pubblici, trasporti e alloggi¹³⁵. Nella fase di utilizzo, la presenza di indicatori – come l’“asterisco”¹³⁶ sull’*Itinerary Card*, segnale di transito in zone a rischio – hanno avuto come conseguenza l’esclusione da determinate zone e obblighi di isolamento, mostrando come un mero indicatore tecnico potesse avere immediate ricadute sul piano giuridico¹³⁷.

La peculiarità del modello cinese di *contact tracing* è da rintracciarsi nella capacità di raccogliere non solo di dati di prossimità e di geolocalizzazione, ma anche nell’obbligatorietà del *check-in* nei luoghi. Difatti, l’adozione – in molte città – di “*venue codes*”, ossia un codice univoco assegnato a un luogo fisico, che venivano esposti all’ingresso di uffici pubblici, ristoranti e complessi residenziali per controllare tutti gli ingressi, ha consentito ricostruzioni rapide di esposizioni “per luogo e per fascia oraria” secondo logiche di correlazione automatizzata tra *database*¹³⁸. In parallelo, l’uso di *big data* per qualificare i singoli come “*close contacts*” e imporre quarantene¹³⁹, con contestazioni sull’accuratezza e sull’opacità dei criteri, hanno offerto

¹³³ Si veda l’ottima ricostruzione di K. Von Carnap-K. Drinhausen-K. Shi-Kupfer, *Tracing. Testing. Tweaking. Approach to data-driven Covid-19 management in China*, Mercator Institute for China Studies (MERICS), 2020.

¹³⁴ National Health Commission of the People’s Republic of China, *Protocol for Prevention and Control of COVID-19*, 2020.

¹³⁵ Per ulteriori approfondimenti, cfr. H. Yu, *Living in the era of codes: a reflection on China’s health code system*, in *Biosocieties*, 2022, 5 ss.

¹³⁶ China Services Info, *Asterisks indicating risky areas removed from travel cards*, 20 giugno 2022.

¹³⁷ F. Liang, *Covid-19 and Health Code: How Digital Platforms Tackle the Pandemic in China*, in *Society and Media*, 2020.

¹³⁸ Cfr. M. Chen-S Xu-L. Husain et al., *Digital health interventions for COVID-19 in China: a retrospective analysis*, in *Intelligence Medicine*, 1, 2021, 29 ss.

¹³⁹ B. Goh, *China rolls out fresh data collection campaign to combat coronavirus*, Reuters, 20 febbraio 2020.

una combinazione di strumenti, con un'altissima capacità di controllo, cui si è accompagnata però una bassa spiegabilità¹⁴⁰. Inoltre, la trasparenza circa i criteri algoritmici adottati per l'assegnazione dei codici colorati è rimasta sostanzialmente opaca¹⁴¹, mentre le vie di ricorso contro decisioni automatizzate sono state, quando presenti, limitate o inefficaci¹⁴². Ne è derivato una trasparenza sostanziale praticamente assente, mentre l'*accountability* è risultata segmentata tra amministrazioni e piattaforme, favorendo la sedimentazione di un'infrastruttura orientata tanto al contenimento epidemiologico quanto al governo ordinario della popolazione¹⁴³. In termini comparativi, l'esperienza cinese mette in luce la distanza tra una trasparenza centrata sul cittadino, quale pretesa di conoscibilità e contestazione, e una trasparenza verticale, funzionale al controllo e alla manutenzione del sistema.

5. Considerazioni conclusive: verso un “costituzionalismo algoritmico dell'emergenza”?

L'analisi comparata dei tre modelli esaminati ha evidenziato una trasformazione strutturale del rapporto tra potere pubblico, infrastrutture digitali e tenuta degli assetti costituzionali. In un arco di tempo brevissimo, sistemi algoritmici e automatizzati – certificazioni, *contact tracing*, modelli previsionali, codici sanitari – sono passati dall'essere mere soluzioni tecniche “di supporto” a meccanismi di allocazione dei diritti, incidendo concretamente su accesso a spazi, lavoro, servizi e circolazione. Tale rapida trasformazione ha reso maggiormente visibile una dinamica già attenzionata dal dibattito sul costituzionalismo digitale, ossia che la configurazione tecnica contribuisce a strutturare le condizioni dell'azione pubblica, diventando un vettore di governo¹⁴⁴. Sicché, le scelte su standard e protocolli, la defi-

¹⁴⁰ Tali aspetti sono stati messi in luce, in particolare, da X. Zhang, *Decoding China's COVID-19 Health Code Apps: The Legal Challenges*, in *HealthCare*, 2022, 6 ss.

¹⁴¹ Per un'analisi accurata circa la regolamentazione degli algoritmi in Cina, v. J. Xu, *Opening the “black box” of algorithms: regulation of algorithms in China*, in *Communication Research and Practice*, 10, 2024.

¹⁴² X. Ding-B. Shang-C. Xie, *et al.*, *Artificial Intelligence in the Covid-19 pandemic: balancing benefits and ethical challenges in China's response*, in *Humanities & Social Sciences Communications*, 2025.

¹⁴³ Cfr. T. Shen-C. Wang, *Regulating China's health code system to prepare for future pandemics*, cit.

¹⁴⁴ Il riferimento classico è all'opera di L. Lessig, secondo cui il codice e l'architettura tecnica funzionano come una forma di regolazione al pari del diritto, del mercato e delle norme sociali, in quanto determinano cosa è possibile fare nello spazio digitale (L. Lessig, *Code: and Other Laws of Cyberspace, Version 2.0*, New York, 2006). Anche altri autori evidenziano il passaggio da una “*governance of infrastructure*” a “*governance by infrastructure*”, in cui il livello di nomi di dominio, indirizzo, protocolli e standard diventa sede di poteri di sorveglianza, censura e di scelte rilevanti sul piano dell'esercizio dei diritti fondamentali (F. Musiani-D.L. Cogburn-L. DeNardis *et al.*, *The Turn to Infrastructure in Internet Governance*, New York, 2016). Infine, l'ampia letteratura sul “*digital constitutionalism*” insiste sul fatto che la governance digitale passa sempre più per standard, protocolli privati, che producono effetti analoghi a norme costituzionali (G. De Gregorio-R. Radu, *Digital constitutionalism in*

nizione delle soglie di rischio, le modalità di trattamento dei dati, la distribuzione delle competenze tra soggetti pubblici e privati e le condizioni per l'*audit* indipendente di un sistema non sono aspetti neutri, ma determinano la misura effettiva delle libertà in contesti ordinari e, con maggiore evidenza, in situazioni emergenziali.

Non a caso, l'analisi comparata restituisce modelli differenziati proprio in ragione della diversa allocazione del potere "tecnico" e della diversa configurazione dei controlli.

Nell'Unione europea, la legittimazione delle infrastrutture sanitarie è stata ricercata nella combinazione tra legalità, standardizzazione tecnica e interoperabilità, rafforzando la *compliance* formale, a scapito però di un'*accountability* sostanziale. Infatti, la scelta procedurale "accelerata" e la non piena istituzionalizzazione di controlli preventivi omogenei – emblematica, in tal senso, la discussione sulla valutazione d'impatto a livello unionale – hanno segnato una distanza incolmabile tra trasparenza come adempimento e trasparenza come condizione di contestabilità effettiva delle decisioni automatizzate.

Negli Stati Uniti, il baricentro è stato opposto, con meno vincoli generali e più regolazione per standard e consorzi tecnici. L'assenza di un'infrastruttura federale unitaria ha favorito rapidità e sperimentazione ma, al contempo, ha accentuato una responsabilità pubblica frammentata e spesso mediata da attori privati, generando deficit di trasparenza sostanziale e strumenti di contestazione non sempre effettivi nella fase acuta della crisi. Qui l'emergenza ha mostrato, con particolare nitidezza, il rischio di "*surveillance laundering*".

In Cina, l'integrazione organica tra strumenti sanitari, amministrazione territoriale e piattaforme digitali ha generato una capacità di intervento senza precedenti, a fronte di un'opacità strutturale incompatibile con l'idea di contestazione effettiva. Il caso "Henan" – in cui i codici sanitari rossi sono stati utilizzati per impedire una mobilitazione cittadina – mostra in modo pragmatico il rischio di "*function creep*", in cui, dinanzi a requisiti tecnici non verificabili, la distinzione tra finalità sanitaria e controllo politico tende a erodersi, soprattutto in ragione del fatto che la possibilità dell'abuso è incorporata nell'architettura stessa.

Alla luce di questo quadro, l'idea di un "costituzionalismo algoritmico dell'emergenza" ha senso solo se non resta formula suggestiva, ma si traduce in criteri operativi in grado di rendere (i) *verificabile* ciò che tende ad essere opaco, attraverso una trasparenza algoritmica non ridotta al codice o alle specifiche, ma estesa a finalità, dati e soglie di rischio che determinano, in concreto, chi accede e chi viene escluso; (ii) *temporaneo* ciò che tende strutturalmente a permanere, mediante *sunset clauses* tecnologicamente in-

the new era of Internet governance, in *International Journal of Law and Information Technology*, 30, 1, 2022, 68 ss.; O. Pollicino, *Regolazione e innovazione tecnologica nell'ordinamento della rete*, in *Rivista AIC*, 2, 2025, 124 ss.).

corporate e controllabili; infine, (iii) *giustiziabile* ciò che è automatico, assicurando rimedi rapidi ed efficaci contro errori e decisioni incomprensibili. Da questi principi e meccanismi dipende, in misura non retorica ma strutturale, la qualità della democrazia nella gestione di nuove emergenze e, più ampiamente, la sostenibilità del progetto democratico-liberale nell'era dell'automazione delle decisioni pubbliche.

Abstract

La pandemia da Covid-19 ha messo a dura prova la tenuta degli ordinamenti costituzionali, trasformandosi in uno “*stress test*” per le garanzie democratiche. In tempi straordinariamente brevi, molti governi hanno incrementato il ricorso a sistemi algoritmici e soluzioni digitali – spesso integrate in infrastrutture proprietarie – la cui opacità ha limitato la trasparenza dei processi e la possibilità di un effettivo controllo democratico. Muovendo da un’analisi comparata, il contributo esamina i modelli adottati nell’Unione europea, negli Stati Uniti e in Cina, in relazione a due strumenti che sono stati utilizzati per contrastare la pandemia: le certificazioni digitali e il *contact tracing*, approfondendo le diverse strategie di *governance* algoritmica messe in campo nei tre contesti e valutando la tenuta degli assetti ordinamentali.

The Covid-19 pandemic has severely tested constitutional systems, turning them into a genuine “*stress test*” for democratic safeguards. Within an extraordinarily short time, many governments intensified their reliance on algorithmic systems and digital solutions – often integrated into proprietary infrastructures – whose opacity has limited procedural transparency and constrained effective democratic oversight. Building on a comparative analysis, this paper examines the models adopted in the European Union, the United States and China in relation to two key tools employed to address the pandemic: digital certificates and contact tracing. It explores the different strategies of algorithmic governance implemented in these three contexts and assesses the resilience of their institutional and legal architectures.

Keywords

pandemia - diritto comparato - trasparenza algoritmica - *accountability* - emergenza

Tra efficienza e tutele: le decisioni amministrative automatizzate in prospettiva comparata*

Angela Ferrari Zumbini, Martina Conticelli,
Marta Infantino

Sommario

1. Introduzione. – 2. Il contesto e la metodologia della ricerca. – 3. La base giuridica per le decisioni amministrative automatizzate. – 4. I settori coinvolti. – 5. I requisiti procedurali. – 6. Casi notevoli in materia di decisioni amministrative automatizzate. – 7. Conclusioni.

1. Introduzione

Il ricorso delle amministrazioni pubbliche (comprese quelle statali, locali, gli enti pubblici e le autorità indipendenti) a processi decisionali automatizzati (*Automated Decision Making* – ADM), più o meno basati su sistemi di intelligenza artificiale (IA), è in aumento quasi ovunque nel mondo¹.

La dipendenza dei progressi tecnologici dalle infrastrutture statali è evidente in Cina. Negli ultimi due decenni il governo cinese ha coltivato la propria industria tecnologica nazionale in stretta collaborazione con il settore privato², rendendo la Cina e gli Stati Uniti i principali concorrenti nella corsa al digitale e all'intelligenza artificiale³. Il governo cinese esporta all'estero molte delle proprie tecnologie digitali, offrendo soprattutto ai paesi emergenti un percorso accessibile verso lo sviluppo digitale e promuovendo le proprie infrastrutture e il proprio modello normativo guidati

* L'articolo è stato sottoposto, in conformità al regolamento della Rivista, a referaggio "a doppio cieco".

Il presente contributo è realizzato nell'ambito del PRIN 2020 "Decision-Making in the Age of Emergencies. New Paradigms in Recognition and Protection of Rights [DeMa]" – CUP C61B22002230001

¹ OECD, *Governing with Artificial Intelligence. The State of Play and Way Forward in Core Government Functions*, 2025; Stanford University, *The AI Index Report 2025* (anche se con riferimento solo alle ADM basate sull'IA). Il concetto di 'stato algoritmico' si riferisce ai molti modi in cui i governi cercano di integrare nel settore pubblico le più recenti innovazioni tecnologiche (infrastrutture digitali, raccolta di Big Data, processi decisionali algoritmici, servizi rivolti agli utenti): Z. Engin – J. Crowcroft – D. Hand – P. Treleaven, *The Algorithmic State Architecture (ASA): An Integrated Framework for AI-Enabled Government*, 2025.

² A. Bradford, *The Global Battle to Regulate Technology*, New York, 2023, 69.

³ Stanford University, *The AI Index Report 2025*, cit., 86 ss.

dallo stato⁴.

Quando si guarda al versante occidentale di queste trasformazioni, si tende spesso a trascurare il ruolo dello stato nello sviluppo e nell'uso delle nuove tecnologie. Il dibattito occidentale degli ultimi anni sulla governance e la regolamentazione algoritmica si è concentrato pressoché interamente sugli attori privati, in particolare sulle grandi imprese tecnologiche americane, considerate i regolatori transnazionali *de facto* dello spazio digitale⁵. Una simile interpretazione riflette accuratamente la posizione nell'ecosistema digitale contemporaneo delle grandi società tecnologiche americane, che includono le MAMAA (Meta, Apple, Microsoft, Amazon, Alphabet) e i molti altri, meno visibili, attori dell'attuale panorama digitale (le società di telecomunicazioni, software e hardware, i fornitori di servizi cloud, le società di analisi dei dati e i broker di dati)⁶. Tuttavia, esaminando più da vicino il modo in cui le imprese statunitensi hanno guidato l'economia digitale, diventa evidente che la crescita del settore tecnologico americano è stata resa possibile grazie agli investimenti pubblici e al forte sostegno del governo, sia a livello nazionale che internazionale⁷. Il coinvolgimento dei poteri pubblici nello sviluppo delle tecnologie di IA negli Stati Uniti, peraltro, è andato ben oltre il semplice finanziamento e sostegno alle aziende tecnologiche: un'indagine del 2020 sull'uso delle tecnologie emergenti nel governo federale ha rivelato che le agenzie federali statunitensi sono all'avanguardia nella sperimentazione dell'IA⁸.

Sebbene per ragioni diverse, anche in Europa il contributo dello stato al finanziamento e allo sviluppo delle nuove tecnologie è stato significativo⁹. In particolare, nei sistemi giuridici di *civil law* prevalenti in Europa, lo stato rimane un fornitore di servizi fondamentale in settori quali la sanità, l'istruzione, l'energia, le comunicazioni e i trasporti¹⁰. Di conseguenza, quasi tutte le sperimentazioni con le ADM in detti settori implicano l'uso della

⁴ A. Bradford, *The Global Battle to Regulate Technology*, cit., 290.

⁵ P. Schiff Berman, *Understanding Global Legal Pluralism: From Local to Global, from Descriptive to Normative*, in P. Schiff Berman (ed.), *The Oxford Handbook of Global Legal Pluralism*, Oxford, 2020, 1 ss.

⁶ Nella ricca letteratura sull'argomento in lingua inglese, si v. S. Zuboff, *The Age of Surveillance Capitalism*, New York, 2019; C. O'Neil, *Weapons of Math Destruction*, New York, 2016; F. Pasquale, *The Black Box Society*, Boston, 2015.

⁷ A. Bradford, *The Global Battle*, cit., 265; National Research Council, *Funding a Revolution: Government Support for Computing Research*, Washington, 1999, 198 ss. Ad esempio, l'Amministrazione per lo sviluppo economico degli Stati Uniti, un ufficio del Dipartimento del Commercio degli Stati Uniti, gestisce attualmente un programma di aiuti multimilionario volto a finanziare i 'Tech Hub', ovvero le regioni del paese che dispongono di beni e risorse con il potenziale per diventare competitive a livello globale nelle tecnologie e nelle industrie del futuro.

⁸ D. Freeman Engstrom – D.E. Ho – C.M. Sharkey – M.-F. Cuéllar, *Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies*, 15.

⁹ Non vi è tuttavia alcun dubbio che l'Europa, e in particolare l'Unione europea (UE), sia attualmente in ritardo nella corsa tecnologica: Stanford University, *The AI Index Report 2025*, cit. 86.

¹⁰ M. Pargendler, *The Role of the State in Contract Law: The Common Law-Civil Law Divide*, in *Yale Journal of International Law*, 43, 2018, 143 ss.; J.H. Merryman – R. Pérez-Perdomo, *The Civil Law Tradition*, Stanford, 2007, 3a ed., 96.

tecnologia da parte delle autorità pubbliche.

In questo contesto, il presente articolo mira a sottolineare il ruolo dello stato non solo come regolatore, ma anche come utente significativo e occasionale sviluppatore di nuove tecnologie, in particolare in Europa e negli Stati Uniti. A tal fine, il lavoro offre una panoramica comparatistica su come le amministrazioni statali, sui due lati dell'Oceano Atlantico, utilizzano softwares e/o sistemi di IA per assumere decisioni amministrative parzialmente o completamente automatizzate. Piuttosto che trattare in generale della digitalizzazione delle infrastrutture pubbliche, l'articolo si concentra sulla misura in cui le amministrazioni possono ricorrere e ricorrono a procedimenti automatizzati (basati o no sull'IA) per l'assunzione di decisioni amministrative. Saranno esaminati in particolare quattro profili fondamentali: il quadro normativo su cui si fonda l'utilizzo di ADM da parte delle pubbliche amministrazioni, i settori nei quali tale utilizzo è più frequente, le garanzie procedurali che devono accompagnarlo e le principali decisioni giurisprudenziali adottate al riguardo. Le fonti per questo studio provengono largamente dai risultati scientifici raccolti nel corso del progetto di ricerca PRIN 2022 intitolato *"The dark side of algorithms under the comparative lens: automated administrative decisions between efficiency and due process"* (AutAD)¹¹.

Più specificamente, dopo alcuni cenni al contesto della ricerca e alla sua metodologia (par. 2), il par. 3 esamina le basi giuridiche per l'adozione di decisioni amministrative automatizzate nelle giurisdizioni oggetto del presente studio, analizzando la gamma di soluzioni adottate per giustificare il ricorso all'ADM nel settore pubblico. Come mostrerà il par. 4, esiste una notevole varietà anche per quanto riguarda i settori nei quali l'e-government e l'amministrazione digitale sono più sviluppati, nonché in termini di utilizzo attuale dell'IA nelle attività amministrative. È interessante notare come le differenze nell'affidamento quotidiano all'ADM sembrano essere correlate non tanto alla base normativa esistente, quanto al quadro giuridico complessivo e al livello di sviluppo tecnologico. Per ciò che concerne i requisiti che l'ADM dovrebbe rispettare – per esempio la protezione dei dati, la trasparenza, la supervisione umana e la contestabilità delle decisioni amministrative –, il par. 5 dimostrerà che sulla carta esiste una divergenza significativa tra i paesi dell'Unione europea (UE) e quelli extra-UE. Tuttavia, l'analisi delle decisioni fino ad ora emesse circa lo stato automatizzato, presentata nel par. 6, metterà in luce come il vero divario tra le giurisdizioni esaminate non sia legato esclusivamente all'architettura normativa di ciascun sistema, ma piuttosto derivi (anche) da fattori giuridici e tecnologici più ampi. Il par. 7 concluderà riassumendo i principali risultati e suggerendo ulteriori aree di ricerca.

¹¹ Il progetto è stato finanziato quale Progetto di ricerca di rilevanza nazionale (PRIN) dal Ministero dell'Università e della Ricerca (sovvenzione n. 2022LSRL82) ed è coordinato a livello nazionale da una delle autrici, Angela Ferrari Zumbini. Sono coinvolte nel progetto tre università: l'Università Federico II di Napoli (unità coordinata da Angela Ferrari Zumbini), l'Università Tor Vergata di Roma (unità coordinata da Martina Conticelli) e l'Università di Trieste (unità coordinata da Marta Infantino).

2. Il contesto e la metodologia della ricerca

Per comprendere meglio la metodologia, la portata geografica e il contenuto della ricerca, è importante collocare il progetto AutAD nel suo più ampio contesto intellettuale.

Come già anticipato, il progetto AutAD si concentra sull'ADM amministrativo. Sebbene esista una crescente letteratura accademica sull'ADM amministrativo, la maggior parte degli studi condotti finora si è però focalizzata su casi e normative nazionali piuttosto che sul confronto multi-giurisdizionale fra più paesi¹². Il progetto AutAD, al contrario, mira a esplorare alcuni sviluppi imposti alle amministrazioni pubbliche dall'attuale transizione tecnologica in una prospettiva comparatistica. Riteniamo che sia cruciale capire come, quando e con quali garanzie le autorità pubbliche nazionali possono utilizzare le tecnologie emergenti, soprattutto considerato il numero limitatissimo di strumenti internazionali di diritto uniforme esistenti in materia¹³. Rispecchiando l'area di competenza del team di ricerca, si è principalmente rivolta l'attenzione a Europa e Stati Uniti, sebbene il progetto abbia coinvolto anche sistemi giuridici dell'Estremo Oriente e sistemi sovranazionali, come la UE.

Per quanto riguarda la metodologia, il progetto si pone in linea di continuità con una precedente e più ampia iniziativa di stampo comparatistico: il *Common Core of European Administrative Laws* (CoCEAL). CoCEAL è un progetto avviato nel 2016 da Giacinto della Cananea e Mauro Bussani per studiare i punti in comune e le divergenze tra i diritti amministrativi dei paesi europei¹⁴. A sua volta il progetto CoCEAL poggiava su una pre-esistente iniziativa europea, molto nota negli ambienti del diritto comparato: il progetto *The Common Core of European Private Law* (*Common Core*). Il *Common Core*, lanciato nel 1993 da Ugo Mattei e Mauro Bussani, ha esaminato e tuttora esamina le convergenze e le differenze tra i sistemi di diritto privato

¹² Esistono tuttavia eccezioni, in particolare in Europa. Si v. per esempio i contributi raccolti nel numero speciale della rivista CERIDAP del 2023, e in particolare lo studio introduttivo di H.C.H. Hofmann, *Comparative Law of Public Automated Decision-Making. An Outline*, in CERIDAP, 2023, 1 ss.; H.C.H. Hofmann – F. Pflücke (eds.), *Governance of Automated Decision-Making and EU Law*, Oxford, 2024; Commissione europea, *AI Watch. Panorama europeo sull'uso dell'intelligenza artificiale da parte del settore pubblico*, 2022; Consiglio d'Europa, *Artificial Intelligence and Administrative Law*, 2022.

¹³ L'unico testo internazionale vincolante finora adottato in questo campo è la Convenzione quadro del Consiglio d'Europa sull'intelligenza artificiale e i diritti umani, la democrazia e lo stato di diritto, aperta alla firma nel settembre 2024, su cui si veda A. Ferrari Zumbini – P. Monaco – S. Venier (a cura di), *The Council of Europe Framework Convention on AI: Comparative, EU, International, and Sectoral Perspectives*, Special Issue di *Italian Journal of Public Law*, 17, 2025. A livello europeo, oltre ovviamente all'AI Act, l'European Law Institute – un think tank accademico che funge da consulente informale dell'Unione europea – ha pubblicato nel 2022 delle norme modello riguardo l'uso di ADM da parte delle autorità pubbliche (European Law Institute, *Model Rules on Impact Assessment of Algorithmic Decision-Making Systems Used by Public Administration*).

¹⁴ Si v. coceal.it/, e G. della Cananea, *The Common Core of European Administrative Laws. Retrospective and Prospective*, Leiden, 2023; G. della Cananea – M. Bussani, *The 'Common Core' of administrative laws in Europe: A framework for analysis*, in *Maastricht Journal of European and Comparative Law*, 26, 2019, 217 ss. I risultati del progetto CoCEAL sono pubblicati in una serie dedicata della Oxford University Press.

delle giurisdizioni europee¹⁵.

In ordine cronologico, il *Common Core* ha adottato e perfezionato una metodologia di ricerca basata sul cosiddetto “approccio fattuale” (sperimentato per la prima volta da Rudolf Schlesinger nel suo studio sulla formazione dei contratti alla Cornell University negli anni '60 del secolo scorso¹⁶) e sulla teoria della “dissociazione dei formanti giuridici” (svilupata da Rodolfo Sacco per distinguere tra fonti del diritto visibili e verbalizzate e fonti meno evidenti e consapevoli¹⁷). Il *Common Core* promuove iniziative comparatistiche collettive in cui un gruppo di ricercatori provenienti da una molteplicità di sistemi giuridici europei collabora allo studio di un determinato argomento in prospettiva comparata. La ricerca parte dalla redazione di un questionario fattuale che propone alcuni scenari pratici ipotetici ma realistici; ai relatori nazionali coinvolti nel gruppo di ricerca viene quindi chiesto di spiegare come tali scenari sarebbero giuridicamente affrontati nei loro sistemi nazionali e di indicare gli eventuali fattori extra-giuridici (come le considerazioni politiche, il contesto sociale o i valori religiosi) che potrebbero influire sul risultato finale. La logica alla base di questo modo di lavorare è quella di garantire che i relatori nazionali possano identificare tutte le circostanze che influenzano il diritto nei propri sistemi, incluse quelle che ufficialmente non giocano alcun ruolo¹⁸. Nel redigere le loro risposte, i relatori nazionali interpretano e veicolano la cultura giuridica dei sistemi che rappresentano. Sebbene ciò possa instillare una certa componente soggettiva nelle risposte, limitandone il valore descrittivo, la pluralità di voci raccolte arricchisce comunque lo studio con opinioni che difficilmente potrebbero essere colte da una ricerca comparatistica realizzata individualmente.

È stata proprio questa metodologia a costituire il fulcro del progetto CoCEAL. Tuttavia, nel trasferire la metodologia del *Common Core* sul terreno del diritto amministrativo, si sono resi necessari alcuni cambiamenti metodologici. Una delle principali novità introdotte dalla metodologia CoCEAL è stata l'inserimento, nel tradizionale questionario fattuale del *Common Core*, di una serie di domande astratte, alle quali i relatori nazionali sono stati invitati a rispondere per descrivere le caratteristiche principali dei propri ordinamenti giuridici. La ragione di questa aggiunta risiedeva in un problema ben noto alla comparazione privatistica: le informazioni di taglio comparatistico sui sistemi amministrativi europei diversi dalla tradizionale triade dei sistemi francese, tedesco e inglese erano talmente limitate da necessitare la raccolta di alcune informazioni preliminari in lingua inglese

¹⁵ Sulla storia e i connotati di questo progetto, si v. M. Bussani, *The Common Core of European Private Law Project Two Decades After: A New Beginning*, in *European Lawyer Journal*, 15, 2015, 9 ss.; M. Bussani – U. Mattei, *The Common Core Approach to European Private Law*, in *Columbia Journal of European Law*, 3, 1997, 339 ss. I volumi risultanti dal progetto sono stati pubblicati prima da [Cambridge University Press](#) e poi da [Intersentia](#).

¹⁶ R.B. Schlesinger (ed.), *Formation of Contracts: A Study of the Common Core of Legal Systems*, New York, 1968, 11.

¹⁷ R. Sacco, *Legal Formants. A Dynamic Approach to Comparative Law. Instalment I*, in *American Journal of Comparative Law*, 39, 1991, 1 ss.

¹⁸ Per una descrizione della metodologia del nucleo comune, si vedano gli autori citati sopra.

sugli ordinamenti meno noti¹⁹.

Sulle spalle di questi giganti, il progetto di ricerca AutAD mira a esplorare le regole e gli standard che si applicano alla pubblica amministrazione in Europa e negli Stati Uniti quando si affida all'ADM. Tre università – l'Università Federico II di Napoli, l'Università Tor Vergata di Roma e l'Università di Trieste – stanno conducendo la ricerca, per un totale di dieci membri permanenti del team. Nel contesto del progetto, sono state organizzate due conferenze internazionali nel 2024 – una a Roma a maggio e una a Trieste a settembre – per riunire i relatori nazionali, far prendere loro dimestichezza con la metodologia e discutere i rispettivi punti di vista e approcci. La conferenza di Roma ha coinvolto i relatori americani, cinesi e dell'Europa occidentale per organizzare il lavoro di ricerca, mentre la conferenza di Trieste ha raccolto i relatori dell'Europa centrale e orientale. Ventitré ordinamenti giuridici hanno preso parte alla ricerca: un primo gruppo comprendeva Austria, Cina, Estonia, Francia, Italia, Germania, Spagna, Paesi Bassi, Regno Unito, Stati Uniti e Unione europea, mentre il secondo gruppo comprendeva Albania, Bulgaria, Croazia, Repubblica Ceca, Ungheria, Lettonia, Lituania, Polonia, Romania, Serbia, Slovenia e Turchia. Ai relatori dell'Europa centrale e orientale è stato chiesto di rispondere a domande astratte e di redigere una relazione nazionale che illustrasse lo stato dell'arte dell'uso della tecnologia da parte delle autorità pubbliche²⁰; inoltre, ai relatori americani, cinesi e dell'Europa occidentale è stato chiesto di rispondere a sette scenari fittizi che coinvolgevano l'ADM pubblico e di spiegare come questi casi sarebbero stati affrontati nei rispettivi ordinamenti giuridici²¹. Molti dei riferimenti alle esperienze di Europa e Stati Uniti nei paragrafi seguenti sono tratti dalle risposte prodotte dai relatori nazionali durante la ricerca.

3. La base giuridica per le decisioni amministrative automatizzate

Il primo aspetto che il progetto ha cercato di comprendere è stata la base giuridica per l'uso dell'ADM da parte delle autorità pubbliche nei paesi esaminati. Il fondamento normativo per il processo di digitalizzazione e automazione è costituito da norme preesistenti che vengono interpretate estensivamente in modo tale da includere pure l'ADM? Oppure il passaggio all'ADM ha richiesto riforme *ad hoc* e l'adozione di nuove norme adatte alle nuove tecnologie? Un ulteriore profilo di interesse risiedeva nel comprendere se la base giuridica del processo di digitalizzazione, laddove

¹⁹ G. della Cananea – M. Bussani, *Il 'Common Core' delle leggi amministrative in Europa*, cit., 240.

²⁰ Le relazioni nazionali sono raccolte nel numero speciale 2/2025 dell'*Italian Journal of Public Law*, a cura di M. Bussani – A. Ferrari Zumbini – M. Infantino, *The Law of the Algorithmic State in Central and Eastern Europe*. Nel presente studio vengono presi in considerazione per la comparazione gli ordinamenti europei e quello statunitense.

²¹ Le relazioni nazionali saranno pubblicate nel numero speciale 1/2026 dell'*Italian Journal of Public Law*, a cura di A. Ferrari Zumbini – M. Conticelli, *Automated Administrative Decisions and Due Process: a Comparative Analysis*, di prossima uscita.

esistente, derivasse da leggi generali o norme settoriali.

A questo proposito, è importante notare come in tutti i paesi oggetto di studio si sia recentemente registrato un significativo rinnovamento delle pubbliche amministrazioni. Dall'analisi comparata emergono due grandi ondate di trasformazione²². La prima è iniziata nei primi anni 2000, quando molti paesi si sono concentrati sullo sviluppo dell'e-government e sulla riorganizzazione della pubblica amministrazione per la gestione e l'uso delle informazioni in formato elettronico. Una seconda ondata è emersa a partire dal 2015, e si è largamente centrata sulla creazione di un governo digitale, intelligente e automatizzato in cui varie tecnologie e strumenti, tra cui l'IA, sono parte integrante del funzionamento della pubblica amministrazione, anche nei processi decisionali. Va tuttavia sottolineato come tali trasformazioni non abbiano necessariamente richiesto cambiamenti al quadro normativo, essendo state spesso attuate tramite strategie e indirizzi governativi e politiche di pianificazione locale. Questo è stato uno dei risultati più chiari emersi dalla ricerca: le scelte strategiche delle autorità pubbliche in merito al tipo di tecnologia su cui fare affidamento, in quali settori e per quali compiti, sono spesso determinate meno dai grandi principi costituzionali-normativi, e più dalle decisioni assunte da tali enti nell'ambito della loro autonomia organizzativa, enti che a loro volta sono sensibili alle consuetudini burocratiche e alle contingenze locali²³.

Tuttavia, è anche vero che le attività amministrative in quasi tutti i sistemi giuridici esaminati sono disciplinate da leggi generali nazionali in materia di procedimento amministrativo, i cui nomi e forme variano notevolmente da un sistema giuridico all'altro²⁴. La ricerca ha finora dimostrato che, in quasi tutti i paesi studiati, queste leggi generali si applicano oggi anche alle questioni relative all'amministrazione digitale. In alcuni sistemi giuridici ciò è stato realizzato semplicemente in via interpretativa. Ad esempio, negli Stati Uniti, le agenzie federali hanno sperimentato l'ADM e l'IA nonostante l'*Administrative Procedure Act* del 1946, che regola le agenzie amministrative del governo federale, non sia mai stato modificato a tal fine²⁵. Esempio diverso è quello dell'Austria, dove l'amministrazione pubblica automatizzata è stata resa possibile da un emendamento dell'*Allgemeines Verwaltungsverfahrensgesetz* (AVG – Legge generale sulla procedura

²² Ciò è in linea con lo sviluppo della ricerca e della letteratura sull'e-government e lo stato digitale: Z. Engin – J. Crowcroft – D. Hand – P. Treleaven, *The Algorithmic State Architecture (ASA)*, cit., 5; E.W. Welch, *Introduction to the Research Handbook on E-Government*, in E.W. Welch (ed), *Research Handbook on E-Government*, Cheltenham, 2021, 1 ss.; si v. anche D.R. Troitiño, *eGovernance as a Future Option*, in D.R. Troitiño (ed), *e-Governance in the European Union*, Cham, 2024, 1 ss.

²³ M. Broussard, *More than a Glitch. Confronting Race, Gender, and Ability Bias in Tech*, Boston, 2023, 76; J. Wolswinkel, *Artificial Intelligence and Administrative Law*, cit., 21; D. Freeman Engstrom *et alii*, *Government by Algorithm*, cit., 6.

²⁴ G. della Cananea – L. Parona, *Administrative Procedure Acts in Europe: An Emerging 'Common Core'?*, in *American Journal of Comparative Law*, 72, 2024, 324 ss.; G. della Cananea, *The Common Core of European Administrative Laws*, cit., 56; G. della Cananea – A. Ferrari Zumbini – O. Pfersmann (eds.), *The Austrian Codification of Administrative Procedure. Diffusion and Oblivion (1920-1970)*, Oxford, 2023.

²⁵ C.M. Sharkey – C. Barrena Hyneman, *US Report*, in *Italian Journal of Public Law*, 18, 2026, di prossima pubblicazione; D. Freeman Engstrom *et alii*, *Government by Algorithm*, cit.

amministrativa) del 1991²⁶. Similmente, i testi normativi fondamentali del diritto amministrativo sono stati ampiamente rivisti in molti altri paesi per tenere esplicitamente conto dell'uso delle tecnologie digitali, dell'ADM e dell'IA. Ad esempio, le norme sull'ADM e sull'IA sono ora esplicitamente incorporate nella legge spagnola sul regime giuridico del settore pubblico (2015)²⁷, nel codice francese sulle relazioni tra il pubblico e l'amministrazione (dal 2016)²⁸, nella legge tedesca sul procedimento amministrativo (dal 2017)²⁹, nella legge estone sul procedimento amministrativo (dal 2018)³⁰, nella legge croata sul procedimento amministrativo generale (dal 2021)³¹, nel codice italiano degli appalti pubblici (2023)³², nella legge lettone sulla responsabilità amministrativa (dal 2025)³³. Altri paesi ancora hanno emanato leggi specificamente volte a regolamentare l'amministrazione digitale. Ad esempio, nel 2023 sia l'Albania che l'Ungheria hanno adottato una legislazione speciale per l'e-government e i servizi digitali³⁴. Nel 2024 la Serbia e la Turchia hanno entrambe presentato progetti di legge sull'IA per regolamentarne l'uso nella pubblica amministrazione, sulla falsariga del Regolamento sull'intelligenza artificiale (più comunemente noto come AI Act³⁵) adottato dalla UE³⁶. Nel 2025, l'Italia ha adottato una legge sull'intelligenza artificiale che, oltre a stabilire alcuni principi generali in materia di ricerca, sperimentazione, sviluppo e adozione di sistemi di IA, disciplina

²⁶ M. Züßner, *Austrian Report*, in *Italian Journal of Public Law*, 18, 2026, di prossima pubblicazione. Sulla storia della legge austriaca sul procedimento amministrativo, si v. G. della Cananea – A. Ferrari Zumbini – O. Pfersmann (eds.), *The Austrian Codification of Administrative Procedure*, cit.

²⁷ Art. 41 della *Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público (LRJSP)*. Oltre ad essere uno dei primi paesi ad aver approvato una riforma sullo stato digitale, nel 2021 la Spagna ha anche adottato una Carta dei diritti digitali, il cui articolo XVI tratta dei diritti digitali dei cittadini nei rapporti con la pubblica amministrazione nazionale: *Carta de Derechos Digitales*.

²⁸ Art. L. 311-3-1 del Codice delle relazioni tra il pubblico e l'amministrazione (*Code des relations entre le public et l'administration*, 2015) modificato dalla *loi* n. 2016-1321 del 7 ottobre 2016 *pour une République Numérique*.

²⁹ § 35a della legge sul procedimento amministrativo (*Verwaltungsverfahrensgesetz*, 1976) modificata nel 2016, in vigore dal 2017.

³⁰ §§ 5, 14, 25, 27, 55 della legge sul procedimento amministrativo (*Haldusmenetluse seadus*, 2001) modificata nel 2017, in vigore dal 2018.

³¹ Artt. 71, 75, 94, 98 della legge sul procedimento amministrativo (*Odluka Zakona o općem upravnom postupku*, 2009) modificata nel 2021, in vigore dal 2022.

³² Art. 30 del Decreto Legislativo 31 maggio 2023, n. 36.

³³ Artt. 162 e 164 della legge sulla responsabilità amministrativa (*Administratīvās atbildības likums*, 2018) modificata nel 2024, in vigore dal 2025.

³⁴ Legge n. 43/2023, Verso l'e-government (*Për qeverisjen elektronike*) in vigore dal 2023, e Legge CIII del 2011 sullo stato digitale e alcune norme per la fornitura di servizi digitali (*évi CIII. törvény a digitális államról és a digitális szolgáltatások nyújtásának egyes szabályairól*, 2023) in vigore dal 2024.

³⁵ Regolamento (UE) 2024/1689 del Parlamento Europeo e del Consiglio, del 13 giugno 2024, sull'intelligenza artificiale.

³⁶ M. Milenković – S. Vračar – M. Pešterić, *The Law of the Algorithmic State in Serbia*, in *Italian Journal of Public Law*, 17, 2025, 768 ss.; P. Çağlayan Aksoy, *The Law of the Algorithmic State: The Case of Türkiye*, in *Italian Journal of Public Law*, 17, 2025, 805 ss.

l'uso dell'IA nella pubblica amministrazione e nell'attività giudiziaria³⁷. Nonostante la mancanza di competenze esplicite sia in materia di pubblica amministrazione che nella tecnologia, nell'ultimo decennio la UE è diventata un regolatore molto importante in questo settore, adottando diversi testi che si sono rivelati fondamentali per regolamentare l'uso pubblico delle tecnologie emergenti negli Stati membri e oltre. Il primo atto legislativo degno di nota a questo proposito è senza dubbio il Regolamento generale sulla protezione dei dati (GDPR).³⁸ L'art. 22 del GDPR prevede che, per sottoporre una persona a un ADM che produce effetti giuridici che la riguardano o la influenzano in modo analogo, sia necessario il consenso esplicito dell'interessato oppure una specifica autorizzazione legislativa. Quando tali requisiti sono soddisfatti, il ricorso alle ADM è legittimo – secondo lo stesso articolo – solo se il titolare del trattamento attua misure adeguate a salvaguardare i diritti, le libertà e gli interessi legittimi dell'interessato, che devono includere il diritto all'intervento umano da parte del titolare del trattamento e la facoltà di contestare la decisione³⁹. In Europa, l'art. 22 del GDPR è stata una delle prime disposizioni a prevedere la possibilità dell'ADM sia nel settore privato che in quello pubblico, ed è infatti stata fonte di ispirazione per molte delle riforme nazionali appena menzionate⁴⁰. Più recentemente, nel 2024, la UE ha adottato l'*Interoperable Europe Act*, con l'intento di promuovere l'interoperabilità del settore pubblico in tutta l'Unione⁴¹, seguito pochi mesi dopo dall'AI Act. L'AI Act non si occupa dello stato automatizzato, ma contiene diverse norme e disposizioni che si applicano anche alle autorità pubbliche. Più specificamente, l'AI Act identifica alcuni usi dell'IA come ad alto rischio e impone agli sviluppatori di sistemi di IA destinati a tali usi di rispettare alcuni obblighi *ex ante* (come la predisposizione di un sistema di gestione dei rischi, la produzione di documentazione tecnica, la previsione di una costante supervisione umana e la realizzazione di una valutazione di conformità)⁴². Quasi tutte le forme di IA ad alto rischio elencate nell'allegato III dell'AI Act (ad esempio, l'IA utilizzata: per gestire il traffico stradale e la fornitura

³⁷ Artt. 14-15 della legge 23 settembre 2025, n. 132.

³⁸ Regolamento (UE) 2016/679 del Parlamento Europeo e del Consiglio, del 27 aprile 2016, relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati.

³⁹ Art. 22(3), GDPR. La Corte di giustizia dell'Unione europea ha recentemente adottato le prime sentenze che chiariscono la portata dell'art. 22: CGUE, C-634/21, *OQ contro Land Hessen (Schufa)* (2023) e CGUE, C-203/22, *CK contro Dun & Bradstreet Austria GmbH e Magistrat der Stadt Wien* (2025). La disposizione è in corso di revisione nel quadro della riforma c.d. *Digital omnibus*.

⁴⁰ Ad esempio, le riforme attuate in Francia, Germania e Italia (si v. le note precedenti) sono state tutte ispirate dall'art. 22 del GDPR. Più in generale, sull'impatto dell'art. 22 del GDPR sul diritto pubblico europeo, S.E. Biber, *Between Humans and Machines: Judicial Interpretation of the Automated Decision-Making Practices in the EU*, in H.C.H. Hofmann – F. Pflücke (eds.), *Governance of Automated Decision-Making and EU Law*, cit., 186 ss.

⁴¹ Regolamento (UE) 2024/903 del Parlamento europeo e del Consiglio, del 13 marzo 2024, che stabilisce misure per un livello elevato di interoperabilità del settore pubblico nell'Unione.

⁴² AI Act, artt. 8–22. Molte di queste disposizioni sono attualmente oggetto di ripensamento nel quadro della riforma c.d. *Digital omnibus*.

di acqua, gas, riscaldamento ed elettricità; per determinare l'ammissione alle scuole; per valutare l'idoneità ai programmi di assistenza sociale; per classificare le chiamate di emergenza; per valutare il rischio che una persona fisica commetta un reato; per esaminare le domande di asilo, visto e permesso di soggiorno e per applicare la legge alle controversie portate in tribunale) comportano l'uso dell'IA nell'ADM amministrativo⁴³. A questo proposito, due disposizioni dell'AI Act sono di particolare rilevanza. L'art. 27 dell'AI Act stabilisce che gli enti pubblici che utilizzano sistemi di IA ad alto rischio devono valutare il potenziale impatto di tali sistemi sui diritti fondamentali⁴⁴. L'art. 86 dell'AI Act prevede che ogni volta che un ente pubblico adotta una decisione che «produc[e] effetti giuridici o in modo analogo incid[e] significativamente su tale persona in un modo che essa ritenga avere un impatto negativo sulla sua salute, sulla sua sicurezza o sui suoi diritti fondamentali» sulla base dei risultati di un sistema di IA ad alto rischio, la persona interessata «ha il diritto di ottenere dal deployer spiegazioni chiare e significative sul ruolo del sistema di IA nella procedura decisionale e sui principali elementi della decisione adottata»⁴⁵. In modo indiretto, l'AI Act chiarisce quindi che le amministrazioni pubbliche europee possono legittimamente ricorrere all'ADM basato sull'IA, a condizione che rispettino determinati requisiti. Esso stabilisce inoltre le condizioni alle quali è consentito ricorrere all'ADM basato sull'IA e fornisce alcune garanzie di base per proteggere i destinatari delle decisioni automatizzate basate sull'IA⁴⁶. Sebbene l'AI Act sia direttamente vincolante per gli enti pubblici solo in Europa, il regolamento potrebbe diventare un modello per la regolamentazione dell'IA ben oltre i confini della UE, ammesso che si inneschi il cosiddetto “effetto Bruxelles”⁴⁷. Come dimostra il caso dei progetti di legge serbi e turchi sopra citati, tale effetto potrebbe già essere in atto⁴⁸.

4. I settori coinvolti

Un altro obiettivo generale del progetto AutAD era valutare in che misura le autorità pubbliche utilizzino l'automazione algoritmica e/o l'IA nelle loro operazioni quotidiane.

Come già osservato nel par. 3, in quasi tutti gli ordinamenti giuridici l'avvento dello stato digitale costituisce la seconda fase di una transizione algoritmica cominciata con l'introduzione di forme di e-government.

⁴³ O.M. Puigpelat, *The AI Act from the Perspective of Administrative Law: Much Ado About Nothing?*, in *European Journal of Risk Regulation*, 2024, 1 ss.; O.M. Puigpelat, *The Impact of the AI Act on Public Authorities and on Administrative Procedures*, in *CERIDAP*, 4, 2023, 238 ss.

⁴⁴ AI Act, art. 27.

⁴⁵ AI Act, art. 86(1).

⁴⁶ O.M. Puigpelat, *The AI Act from the Perspective of Administrative Law*, cit., 8.

⁴⁷ A. Bradford, *The Brussels Effect: How the European Union Rules the World*, New York, 2019.

⁴⁸ Si veda anche, in Brasile, il *Projeto de Lei* n° 2338, de 2023, e, in Cile, il *Proyecto de ley* N° de *Boletín* 16821-19.

Mentre l'esperienza dell'e-government è iniziata pressoché ovunque con la creazione di interfacce online e l'utilizzo da parte delle amministrazioni di mezzi di comunicazione elettronici, sia fra loro che nei rapporti con l'esterno, la digitalizzazione attuale incide sempre più sulla sfera individuale delegando determinati compiti amministrativi a sistemi algoritmici intelligenti che agiscono con vari gradi di autonomia.

La portata e la velocità di questa transizione dall'e-government allo stato digitale variano notevolmente da un paese all'altro. Sebbene gli enti pubblici in tutti gli ordinamenti giuridici considerati stiano gradualmente adottando meccanismi di automazione algoritmica, il grado e la forma di tale automazione differiscono in modo significativo. Al riguardo, non sorprende scoprire che gli ordinamenti giuridici con infrastrutture digitali tecnologicamente più sviluppate abbiano consentito l'utilizzo dell'ADM nel settore amministrativo in modo più rapido e diffuso rispetto ai paesi con ecosistemi digitali meno avanzati⁴⁹.

Tra i paesi che hanno compiuto progressi significativi nell'integrazione dell'IA nella pubblica amministrazione, figurano anzitutto gli Stati Uniti. Il governo federale statunitense sta ampliando notevolmente l'uso dell'ADM (per lo più basata sull'IA) per supportare diverse funzioni amministrative. L'ADM basata sull'IA è ora utilizzata in tutte le aree di intervento del governo, compresi il welfare, l'ambiente, il lavoro, l'occupazione e l'edilizia abitativa. Tuttavia, le tre principali aree di utilizzo per settore sono le sanzioni amministrative, la sanità e la regolamentazione finanziaria⁵⁰. Ad oggi, la fonte più completa che descrive in dettaglio l'uso dell'IA da parte del governo federale è il rapporto ACUS 2020, *Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies*, che include un'indagine su 157 casi di utilizzo dell'IA in 142 agenzie federali⁵¹.

In Europa, simili sviluppi sono evidenti negli Stati baltici. L'Estonia è uno dei pochi paesi (e l'estone una delle poche lingue) ad avere una parola speciale per indicare l'IA: "*kratt*". Il termine deriva dal folklore estone e indica una creatura mitologica che può aiutare le persone, ma che necessita comunque di supervisione⁵². Da questa parola è stato coniato il termine "*bürokratt*" per indicare l'IA amministrativa. Il governo estone ha l'obiettivo esplicito di creare uno stato erogatore di servizi personalizzati basati sull'IA, il che significa che i servizi pubblici per i cittadini sono in gran parte automatizzati e che il contatto con le autorità è necessario solo quando il cittadino *non desidera* un servizio regolare (ad esempio, quando non vuole andare in pensione al raggiungimento dell'età pensionabile)⁵³. In Lituania, gli enti pubblici si affidano a sistemi automatizzati per l'amministrazione

⁴⁹ È sufficiente confrontare quanto segue con i dati sugli sviluppi tecnologici contenuti nel rapporto dell'Università di Stanford, *The AI Index Report 2025*, cit.

⁵⁰ C.M. Sharkey – C. Barrera Hyneman, *US Report*, cit.; C.M. Sharkey, *Algorithmic Tools in Retrospective Review of Agency Rules*, 2023.

⁵¹ D. Freeman Engstrom *et alii*, *Government by Algorithm*, cit.

⁵² K. Nymann Metcalf, *Estonian Report*, in *Italian Journal of Public Law*, 18, 2026, di prossima pubblicazione.

⁵³ Un elenco completo dei progetti di IA realizzati nel settore pubblico estone è disponibile online con una breve panoramica della tecnologia utilizzata, dell'obiettivo e dell'area del progetto e altre informazioni rilevanti: koodivaramu.eesti.ee/explore/projects.

fiscale, i procedimenti giudiziari e gli ordini amministrativi. In particolare, il sistema giudiziario lituano è stato digitalizzato per la gestione dei nuovi casi e l'automazione è utilizzata anche per gestire reati minori come le violazioni delle regole del traffico⁵⁴. Anche in Lettonia l'automazione è molto diffusa nella gestione del traffico veicolare e nella riscossione delle imposte: i sistemi automatizzati sono ampiamente impiegati per emettere multe rispetto alle infrazioni stradali e gestire questioni fiscali, con l'obiettivo di ridurre il coinvolgimento umano nelle attività amministrative di routine⁵⁵. Mentre la maggior parte dei paesi dell'Europa occidentale utilizza algoritmi di apprendimento automatico per analizzare i dati al fine di prevenire le frodi fiscali, due paesi, la Francia e i Paesi Bassi, fanno ampio ricorso all'automazione algoritmica. Tra i numerosi esempi che si possono trovare in Francia, si può citare l'uso dell'ADM per determinare l'idoneità dei disoccupati a ricevere sussidi, nonché l'importo e la durata di questi ultimi (France Travail), per velocizzare l'attraversamento delle frontiere in alcuni aeroporti (Parafe) e per assegnare gli studenti alle università al termine della scuola dell'obbligo (Parcoursup)⁵⁶. Un discorso analogo vale per le amministrazioni olandesi, che, a partire dagli anni '80 del secolo scorso, hanno utilizzato sempre più spesso sistemi informatici entro i processi decisionali⁵⁷. Oggi gli enti pubblici olandesi utilizzano ampiamente algoritmi per individuare precocemente possibili rischi, come il rischio di abbandono scolastico precoce, di frode e di crollo strutturale degli edifici, così da consentire alle autorità, compresa la polizia, di allocare le risorse in modo più efficace e di sottoporre persone o gruppi a controlli e indagini ulteriori. Le autorità amministrative ricorrono a procedimenti automatizzati anche per supportare i procedimenti interni e valutare le richieste di sussidi, permessi e prestazioni⁵⁸.

Nell'Europa orientale, merita di essere segnalato il caso della Polonia, dove l'automazione è utilizzata in diversi settori, tra cui la gestione dei rifiuti, il controllo del traffico, lo sviluppo delle città intelligenti e la protezione dei consumatori⁵⁹. Al di fuori della UE, la Turchia ha integrato l'IA in una serie di settori, tra cui la gestione del traffico, le relazioni pubbliche (tramite chatbot) e la sicurezza pubblica. L'Ufficio turco per la trasformazione digitale (DTO) supervisiona l'implementazione dell'IA in questi settori e registra un notevole slancio verso l'utilizzo dell'automazione nelle opera-

⁵⁴ J. Paužaitė-Kulvinskienė – G. Strikaitė-Latušinskaja, *Regulating Automation: The Legal Landscape of 'Automated Administrative Orders' in Lithuania*, in *Italian Journal of Public Law*, 17, 2025, 664 ss.

⁵⁵ E. Danovskis, *The Use of Automated Decision-Making Systems by the Government in Latvia*, in *Italian Journal of Public Law*, 17, 2025, 652 ss.

⁵⁶ M. Lanna, *French Report*, in *Italian Journal of Public Law*, 18, 2026, di prossima pubblicazione.

⁵⁷ U. Bisgaard Ulsrod Røhl, *Automated Decision-making and Good Administration: Views from Inside the Government Machinery*, in *Government Information Quarterly*, 40, 2023.

⁵⁸ L. Verboeket – J. van den Brink, *Dutch Report*, in *Italian Journal of Public Law*, 18, 2026, di prossima pubblicazione.

⁵⁹ P. Tereszkievicz – M. Namysłowska – S. Skalski – A. Olbryk, *Algorithmic Poland: Transforming Public Administration*, in *Italian Journal of Public Law*, 17, 2025, 696 ss.

zioni governative di routine⁶⁰.

La situazione è diversa in altri ordinamenti giuridici, in particolare nell'Europa sud-orientale, dove l'uso dell'ADM da parte degli enti pubblici nella pratica quotidiana è ancora limitato. In Romania, ad esempio, alcuni aspetti degli appalti pubblici e dei trasporti sono stati digitalizzati, ma l'automazione nelle attività amministrative di routine rimane limitata⁶¹. Sebbene l'Albania abbia implementato l'ADM e l'IA nei settori della sicurezza pubblica, del controllo dell'immigrazione, delle relazioni con i cittadini e degli appalti pubblici (con la creazione del primo ministro virtuale, creato e gestito con IA), la pubblica amministrazione nel paese rimane in gran parte manuale e l'uso dell'automazione algoritmica nei processi decisionali comuni è ancora in una fase embrionale⁶². Allo stesso modo, mentre in Serbia l'ADM è utilizzata in alcuni settori quali la sanità, la gestione fiscale e l'immigrazione, il paese è ancora nelle prime fasi di adozione dell'automazione nella pubblica amministrazione⁶³.

Se si considerano i settori più interessati dalla digitalizzazione e dall'automazione, l'ADM, basata o meno sull'IA, si è diffusa sia nell'ambito delle tradizionali funzioni amministrative – come l'identificazione dei cittadini, la sicurezza, la polizia, l'immigrazione e l'amministrazione fiscale –, che nella fornitura di servizi pubblici – come i trasporti, il welfare, la sanità e, più in generale, gli appalti pubblici. In molti paesi, gli algoritmi svolgono anche funzioni amministrative. Ad esempio, in Italia, sistemi algoritmici sono utilizzati per assegnare i posti a chi (come gli insegnanti) abbia con successo superato un concorso pubblico⁶⁴. In Polonia, l'agenzia per la protezione dei consumatori ha sviluppato uno strumento basato sull'IA chiamato Arbuz, che determina la probabilità che una clausola contrattuale sia illegale. Arbuz esegue un'analisi preliminare dei documenti per identificare le disposizioni dei termini contrattuali standard che potrebbero essere vessatori in quanto gravemente lesivi dei diritti dei consumatori. Le raccomandazioni generate dall'algoritmo possono poi essere verificate da un responsabile umano del caso che prende la decisione finale sull'opportunità di considerare o no una clausola abusiva⁶⁵.

5. I requisiti procedurali

Comprendere le condizioni alle quali le amministrazioni pubbliche possono ricorrere all'ADM è un obiettivo fondamentale del progetto AutAD.

⁶⁰ P. Çağlayan Aksoy, *The Law of the Algorithmic State*, cit.

⁶¹ S. Lodoabă – C. Rus – D.C. Dragos, *The Algorithmic Race – Insights from the Romanian Public Sector*, in *Italian Journal of Public Law*, 17, 2025, 737 ss.

⁶² E. Methasani Çani – A. Mazellu, *Navigating the Algorithmic Shift and the Legal Implications of AI in Albania's Public Administration*, in *Italian Journal of Public Law*, 17, 2025, 497 ss.

⁶³ M. Milenković – S. Vračar – M. Pešterić, *The Law of the Algorithmic State in Serbia*, cit.

⁶⁴ D.U. Galetta – S. D'Ancona, *Italian Report*, in *Italian Journal of Public Law*, 18, 2026, di prossima pubblicazione.

⁶⁵ P. Tereszkievicz – M. Namysłowska – Szymon Skalski – Aleksandra Olbryk, *Algorithmic Poland*, cit.

Nel corso del progetto, è stato chiesto ai relatori nazionali di spiegare i requisiti procedurali che, nel loro paese, si applicano all'ADM per tutelare i diritti dei cittadini e garantire lo stato di diritto. In particolare, i relatori nazionali sono stati chiamati a chiarire se e in che misura, entro il loro sistema giuridico di provenienza, le procedure amministrative automatizzate debbano garantire la protezione dei dati, la trasparenza, il costante controllo umano e il diritto di contestare le decisioni assunte. Va da sé che l'importanza di tali requisiti procedurali dipende dalla forma dell'azione amministrativa in questione, dalla fase della procedura in cui avviene l'automazione e dal livello di democraticità delle amministrazioni. Più un'amministrazione è autoritaria, maggiore è la crucialità delle garanzie individuali definite dal diritto amministrativo.

Non sorprende che i requisiti per l'ADM siano progettati, concepiti e regolamentati in modo diverso nelle diverse giurisdizioni, offrendo ai cittadini livelli di protezione variabili. Come dimostra l'analisi che segue, dedicata separatamente a ciascuno dei requisiti sopra menzionati, esiste un ampio divario tra i paesi della UE e quelli extra-UE, soprattutto per quanto riguarda la protezione dei dati e la supervisione umana. Almeno sulla carta, le garanzie offerte dal diritto dell'Unione europea sono significativamente più numerose e rigorose di quelle offerte altrove.

5.1. La protezione dei dati

Per quanto riguarda la privacy e la protezione dei dati, si può tracciare una netta distinzione tra i paesi della UE e quelli extra-UE. In tutti gli stati membri della UE, oltre che nel Regno Unito, i requisiti in materia di privacy sono rigorosamente regolamentati dal GDPR per quanto riguarda i dati personali. I sistemi di IA che trattano dati personali devono rispettare il GDPR, quali i principi di trattamento lecito, equo e trasparente dei dati, di limitazione delle finalità e di minimizzazione dei dati raccolti⁶⁶. Inoltre, l'art. 22 del GDPR vieta di prendere decisioni basate esclusivamente sul trattamento automatizzato dei dati personali, a meno che l'interessato non abbia espressamente acconsentito al trattamento o non vi sia una chiara base giuridica per farlo; lo stesso articolo prevede anche il diritto degli interessati di contestare la decisione assunta e di richiedere l'intervento di un essere umano⁶⁷.

Sebbene il GDPR sia vincolante solo per le amministrazioni pubbliche degli Stati membri della UE, occorre notare come il prestigio del modello UE sul punto, talvolta legato all'aspirazione di divenire membri dell'Unione, abbia portato molti ordinamenti giuridici estranei alla UE ad adottare una legislazione simile al GDPR. Ciò vale in particolare per l'Albania e la Turchia, ma anche per molti stati americani⁶⁸.

⁶⁶ Art. 5 GDPR.

⁶⁷ Art. 22 GDPR, sul quale si veda anche *supra*, par. 3.

⁶⁸ Cfr. E. Methasani Çani – A. Mazellu, *Navigating the Algorithmic Shift*, cit. (Albania); P. Çağlayan Aksoy, *The Law of the Algorithmic State*, cit. (Turchia); C.M. Sharkey – C. Barrena Hyneman, *US Report*, cit.

5.2. La trasparenza

Nella maggior parte dei paesi investigati, gli obblighi di trasparenza sono un requisito procedurale fondamentale dei procedimenti amministrativi. In teoria, il requisito di trasparenza dovrebbe estendersi anche alle decisioni amministrative automatizzate; tuttavia, nella pratica, i sistemi giuridici esaminati presentano una notevole varietà di soluzioni e livelli di protezione.

Da un lato, alcuni paesi richiedono una trasparenza quasi assoluta nell'assunzione delle decisioni automatizzate. Ad esempio, ai sensi del Codice francese sulle relazioni tra il pubblico e l'amministrazione, un'amministrazione pubblica che si avvale di mezzi automatizzati per assumere una decisione riguardo a una persona ha ora quattro obblighi: fornire informazioni generali sull'uso degli algoritmi, fare esplicito riferimento alle finalità del trattamento, offrire ragguagli sulla decisione automatizzata a richiesta della parte interessata e, infine, divulgare il codice sorgente dell'algoritmo⁶⁹.

In Spagna, l'art. 11.1.i del Regio Decreto 203/2021, oltre a imporre agli organi della pubblica amministrazione l'obbligo di fornire informazioni relative alle azioni amministrative a disposizione delle parti interessate, stabilisce che la pubblica amministrazione debba rendere noto al pubblico l'elenco delle azioni automatizzate, insieme alla loro configurazione e al processo di apprendimento sottostante i sistemi di IA. Requisiti simili sono previsti a livello regionale: basti considerare, a titolo esemplificativo, i casi delle Comunità della Catalogna, della Galizia e di Valencia⁷⁰.

Allo stesso modo, la legislazione lettone che consente l'uso dell'ADM nei settori fiscali stabilisce esplicitamente che le persone debbano essere informate quando le decisioni che le riguardano sono state assunte da sistemi automatizzati e debbano poter avere accesso alla spiegazione circa il perché tali decisioni sono state raggiunte⁷¹.

Altri paesi, pur avendo compiuto sforzi notevoli per migliorare la trasparenza nell'ADM, non si sono ancora dotati di quadri specifici per lo sviluppo tecnologico. Sebbene l'art. 22 del GDPR (quando sono coinvolti dati personali) e l'art. 86 dell'AI Act (quando sono coinvolti sistemi ad alto rischio) stabiliscano in seno alla UE che gli interessati hanno il diritto di essere informati sulla logica alla base delle decisioni automatizzate e sulle spiegazioni delle decisioni basate sul trattamento automatizzato, la maggior parte dei paesi dell'Europa centrale e orientale sembra essere tuttora nelle fasi iniziali della creazione di quadri normativi volti a imporre misure di trasparenza e divulgazione pubblica minima degli strumenti algoritmici e di IA nel settore pubblico⁷².

⁶⁹ Artt. L.300-2, L.300-3, L.311-1-3, L.311-3-1, L.312-1-3 del Codice sulle relazioni tra il pubblico e l'amministrazione; si v. anche M. Lanna, *French Report*, cit.

⁷⁰ Per quanto riguarda la Catalogna, v. l'art. 54.2 del Decreto 76/2020 (4 agosto 2020); per quanto riguarda la Galizia, v. la Legge 4/2019 (17 luglio 2019); per quanto riguarda Valencia, v. la Legge 1/2022 (13 aprile 2022).

⁷¹ E. Danovskis, *The Use of Automated Decision-Making Systems*, cit.

⁷² A. Ferrari Zumbini – M. Conticelli, *The Law of the Algorithmic State in Central and Eastern Europe. Concluding Remarks*, in *Italian Journal of Public Law*, 17, 2025, 850 ss.

5.3. Il controllo umano

Come appena ricordato, negli stati membri della UE, l'art. 22 del GDPR (quando sono coinvolti dati personali) e l'art. 86 dell'AI Act (quando sono coinvolti sistemi ad alto rischio) svolgono un ruolo cruciale nel garantire che i cittadini abbiano il diritto di richiedere l'intervento umano nell'ADM amministrativo. Ciò significa che, entro la UE, le decisioni automatizzate con effetti giuridici o un impatto altrettanto significativo devono essere sempre soggette, o quanto meno assoggettabili, alla supervisione umana. Inoltre, paesi come la Lettonia, la Lituania e la Polonia dispongono di quadri normativi ben sviluppati che impongono il coinvolgimento umano nell'assunzione di decisioni automatizzate, in particolare per quanto riguarda i servizi fiscali e giudiziali⁷³.

I paesi al di fuori della UE non sembrano avere disposizioni giuridiche specifiche che richiedono il coinvolgimento umano nell'ADM amministrativo, anche se si può ragionevolmente supporre che i paesi più vicini all'Unione modificheranno presto le loro regole sul punto per allineare il diritto nazionale al GDPR e all'AI Act⁷⁴.

5.4. Il diritto di contestare le decisioni

Il diritto di contestare le decisioni amministrative automatizzate è strettamente connesso al diritto di richiedere l'intervento umano. A questo proposito, la maggior parte dei paesi della UE dispone di solidi quadri normativi a livello nazionale, con regole che definiscono chiaramente in che modo i cittadini possono contestare le decisioni automatizzate e chiederne la revisione, sia in generale che in settori specifici, come la fiscalità e la gestione del traffico⁷⁵.

Al di fuori della UE, i paesi europei confinanti come l'Albania, la Serbia e la Turchia dispongono attualmente di meccanismi generali per la contestazione delle decisioni amministrative, i quali probabilmente potrebbero applicarsi anche per analogia alle decisioni automatizzate. È tuttavia importante sottolineare come questi paesi stiano tutti cercando di allineare maggiormente il loro quadro giuridico agli standard adottati in seno alla UE⁷⁶. Al contrario, nel Regno Unito e negli Stati Uniti esistono solo rimedi generali. Più specificamente, nel Regno Unito le norme in materia di contestazione delle decisioni amministrative sono dettate sia dalla legislazione sul punto (come applicata dai tribunali) che dai principi del *common law*. I principi normativi elaborati dal governo, sotto forma di strumenti di *soft*

⁷³ Cfr. E. Danovskis, *The Use of Automated Decision-Making Systems*, cit. (Lettonia); J. Pauzaitė-Kulvinskienė – G. Strikaitė-Latušinskaja, *Regulating Automation*, cit. (Lituania); P. Tereszkiewicz – M. Namysłowska – S. Skalski – A. Olbryk, *Algorithmic Poland*, cit. (Polonia).

⁷⁴ Cfr. E. Methasani Çani – A. Mazellu, *Navigating the Algorithmic Shift*, cit. (Albania); M. Milenković – S. Vračar – M. Pešterić, *The Law of the Algorithmic State in Serbia*, cit. (Serbia); P. Çağlayan Aksoy, *The Law of the Algorithmic State*, cit. (Turchia).

⁷⁵ A. Ferrari Zumbini – M. Conticelli, *The Law of the Algorithmic State*, cit.

⁷⁶ Si v. gli autori citati sopra.

law (principalmente il *White Paper 2023. A pro-innovation approach to AI regulation*), non aggiungono molto ai principi generali di *common law* applicabili al ricorso e al controllo giurisdizionale. Tuttavia, essi richiedono agli sviluppatori e agli utilizzatori di IA (nell'ambito delle competenze delle autorità di regolamentazione) di fornire chiarezza agli utenti su come contestare i risultati o le decisioni dell'IA. Negli Stati Uniti, l'Administrative Procedure Act prevede che sia sempre possibile attivare il controllo giurisdizionale rispetto all'azione di un'agenzia pubblica laddove quest'ultima abbia assunto una decisione «*arbitrary, capricious, an abuse of discretion, or otherwise not in accordance with law*»⁷⁷.

6. Casi notevoli in materia di decisioni amministrative automatizzate

Poiché la giurisprudenza è una fonte fondamentale di norme per lo stato automatizzato sia nelle giurisdizioni di *common law* che in quelle di *civil law*, nell'ambito del progetto AutAD si è chiesto ai relatori nazionali di esaminare non solo la legislazione esistente e futura, ma anche le sentenze giudiziali emesse con riguardo a casi di contestazione di decisioni amministrative automatizzate.

A questo proposito, la ricerca ha evidenziato una frattura che sembra avere poco a che fare con gli assetti normativi esaminati nei parr. 3 e 5 e pare per lo più legata al più ampio contesto giuridico e livello di sviluppo tecnologico proprio a ciascun paese. La linea principale di divisione al riguardo corre fra i paesi dell'Europa centro-orientale da un lato e gli Stati Uniti e i paesi dell'Europa occidentale dall'altro lato. Il numero di controversie intentate contro lo stato algoritmico negli Stati Uniti è aumentato costantemente a partire dal 2010; in Europa occidentale un fenomeno analogo è riscontrabile dal 2015 circa. Questa giurisprudenza in espansione mostra come il passaggio all'automazione e ai sistemi intelligenti è un processo di tentativi ed errori che spesso comporta costi notevoli per coloro che vi sono coinvolti, e in particolare per i settori più vulnerabili della popolazione, come le persone con disabilità, i disoccupati, i beneficiari di assistenza sociale e gli immigrati. Al contrario, i contenziosi derivanti dall'uso pubblico dell'ADM sono, per il momento, praticamente inesistenti nell'Europa centro-orientale. Ciò può essere dovuto all'introduzione relativamente recente dell'IA nella pubblica amministrazione e/o alla mancanza di quadri giuridici dedicati alle innovazioni tecnologiche, ma può anche essere legato a tassi generalmente ridotti di contenzioso amministrativo⁷⁸.

Più nel dettaglio, l'analisi delle tendenze del contenzioso negli Stati Uniti mostra che le città, i dipartimenti e le agenzie pubbliche americane ricor-

⁷⁷ Si v. l'Administrative Procedure Act, 5 U.S.C. § 553(b); si v. anche C.M. Sharkey – C. Barrena Hyneman, *US Report*, cit.

⁷⁸ Molti dei paesi dell'Europa centrale e orientale appartengono alla cosiddetta famiglia giuridica post-socialista, che storicamente è stata caratterizzata da bassi tassi di contenzioso, soprattutto se confrontati con quelli dell'Europa occidentale e degli Stati Uniti: A. Uzelac, *Survival of the Third Legal Tradition?*, in *Supreme Court Law Review*, 49, 2010, 377 ss.

rono ad algoritmi per valutare le prestazioni degli insegnanti e risolvere i loro contratti⁷⁹, per indagare sulle frodi nei sussidi di disoccupazione⁸⁰, per erogare prestazioni di invalidità⁸¹, e per effettuare una videosorveglianza urbana costante tramite droni⁸². In uno di questi casi, che merita menzione per la sua enormità, lo stato del Michigan aveva investito 47 milioni di dollari in un contratto privato per sviluppare un algoritmo in grado di imparare a individuare le frodi nei sussidi di disoccupazione sulla base dei dati passati. Il “*Michigan Integrated Data Automated System*” (MIDAS) è stato utilizzato dallo stato del Michigan fra il 2013 e il 2015 per accusare di frode migliaia di cittadini e revocare loro le indennità fino a quel momento ricevute. Alcuni anni dopo, i controlli effettuati dai contabili governativi hanno dimostrato che il 93% degli accertamenti di frode rilevati da MIDAS erano falsi positivi⁸³.

Nel Regno Unito, gli esperimenti con l’ADM finiti in tribunale includono il tentativo della polizia gallese di utilizzare telecamere di sorveglianza e strumenti di riconoscimento facciale⁸⁴, l’utilizzo da parte delle poste del regno di un software di contabilità difettoso per la gestione delle filiali postali⁸⁵, il calcolo automatizzato degli aiuti economici per i lavoratori autonomi durante la pandemia di COVID-19 e lo spionaggio algoritmico dell’intera popolazione ad opera di varie agenzie di intelligence britanniche⁸⁶.

Anche sul continente, i contenziosi stanno aumentando nelle giurisdizioni dell’Europa occidentale. I tribunali francesi hanno già trattato un nume-

⁷⁹ *Hous. Fed’n of Teachers Local 2415 v. Hous. Indep. Sch. Dist.*, 251 F. Supp. 3d 1168 (S.D. Tex. 2017) (sull’algoritmo privato utilizzato dal distretto scolastico indipendente di Houston per valutare le prestazioni degli insegnanti).

⁸⁰ Cfr. ad esempio *Scott v. Department of Labor and Economic Opportunity*, 25 maggio 2023 (2023 Mich. App. LEXIS 3755) (sul programma automatizzato di rilevamento delle frodi appaltato privatamente dal Michigan).

⁸¹ *K.W. v. Armstrong*, 28 marzo 2016 (180 F. Supp. 3d 703) (sul programma di fattura privata adottato in Idaho per valutare le prestazioni di invalidità Medicaid; l’azione collettiva è ancora in corso); *Michael T. v. Crouch*, 26 marzo 2018 (2018 U.S. Dist. LEXIS 49598) (sullo strumento adottato in West Virginia per la valutazione della corretta erogazione delle prestazioni pubbliche dovute in caso di invalidità).

⁸² *Leaders of a Beautiful Struggle v. Balt. Police Dep’t*, 24 giugno 2021 (2F.4th 330) (sul servizio di sorveglianza tramite droni fornito da una società commerciale al Dipartimento di Polizia di Baltimora).

⁸³ Cfr. ad esempio *Zynda v. Arwood*, 175 F. Supp. 3d 791 (E.D. Mich. 2016); *Scott v. Department of Labor and Economic Opportunity*, 25 maggio 2023 (2023 Mich. App. LEXIS 3755). Si v. anche S.M. Gipson Rankin, *The Midas Touch: Atuahene’s ‘Stategraft’ and Unregulated Artificial Intelligence*, in *New York University Law Review Online*, 98, 2023, 225 ss.

⁸⁴ *R (Bridges) v. The Chief Constable of South Wales Police* (2020) EWCA Civ 1058 (sulla legittimità dell’uso di telecamere di sorveglianza e strumenti di riconoscimento facciale, nonostante alcune inaccurately nella redazione del documento di valutazione dei rischi ai sensi del GDPR).

⁸⁵ *Alan Bates and Others v Post Office Limited* (2019) EWHC 3408 (QB) (sui difetti dell’algoritmo utilizzato dal Post Office).

⁸⁶ Nel complesso, l’acquisizione e il trattamento di dati di massa sono stati ritenuti legittimi sia dalla Corte europea dei diritti dell’uomo che dalla Corte di giustizia dell’Unione europea: CEDU, *Big Brother Watch v. Regno Unito*, 25 maggio 2021, ricorsi n. 58170/13, 62322/14 e 24960/15; CGUE, C-623/17, *Privacy International v. Segretario di Stato per gli Affari Esteri e del Commonwealth e altri*, 6 ottobre (2020).

ro considerevole di contestazioni riguardanti ad esempio l'uso di sistemi di riconoscimento facciale biometrico da parte delle scuole nei quartieri svantaggiati⁸⁷, l'istituzione di un software segreto a livello nazionale per l'abbinamento degli studenti alle università⁸⁸, l'utilizzo del riconoscimento facciale per accedere a un'app sanitaria nazionale⁸⁹, e, durante la pandemia di COVID-19, la creazione di un *Health Data Hub* che raccoglieva le informazioni mediche della popolazione e le memorizzava su server situati in Francia, Paesi Bassi e Irlanda⁹⁰. I tribunali olandesi sono intervenuti per pronunciarsi sulla legittimità dell'uso da parte del governo di software per concedere autorizzazioni ambientali⁹¹, valutare il valore dei terreni a fini fiscali⁹² e prevedere il rischio di frodi previdenziali da parte dei beneficiari di assistenza sociale⁹³.

Un sistema algoritmico destinato a identificare i gruppi vulnerabili aventi diritto a bollette elettriche scontate è stato contestato (senza successo) dinanzi ai tribunali spagnoli⁹⁴. In Italia, il software utilizzato dal Ministero dell'Istruzione per coprire i posti vacanti di insegnanti a livello nazionale ha suscitato ripetute lamentele, la maggior parte delle quali sono state in gran parte accolte⁹⁵. Altrettanto successo ha avuto in Austria la causa giudiziaria promossa contro l'algoritmo *Arbeitsmarktservice*, progettato per

⁸⁷ Tribunale di Marsiglia, 27 febbraio 2020, n. 1901249 (per il quale il sistema non era conforme al GDPR).

⁸⁸ Corte costituzionale (*Conseil constitutionnel*), 3 aprile 2020, n. 2020-834 QPC (sulla legittimità dell'uso e della segretezza del software). Nel 2018, la stessa corte aveva emesso un'altra sentenza, affermando che, in generale, l'amministrazione statale possa legittimamente avvalersi di ADM e IA, purché rispetti le norme di legge: Corte costituzionale (*Conseil constitutionnel*), 12 ottobre 2018, n. 2018-765 QPC.

⁸⁹ Consiglio di Stato (*Conseil d'État*), 4 novembre 2020, n. 432656 (che ha confermato la legittimità dell'uso del riconoscimento facciale nell'app, a condizione che l'app non fosse obbligatoria e che esistessero altri mezzi per effettuare il login).

⁹⁰ Consiglio di Stato (*Conseil d'État*), 19 giugno 2020, n. 440916 (che ha confermato la legittimità della banca dati).

⁹¹ Consiglio di Stato (*Raad van State*), 17 maggio 2017, ECLI:NL:RVS:2017:1259 (cosiddetto *Aerius I*) (in cui si è sostenuto che la pubblica amministrazione non possa fare affidamento su sistemi algoritmici i cui contenuti non sono accessibili ai cittadini e alla stessa pubblica amministrazione).

⁹² Corte Suprema (*Hoge Raad*), 17 agosto 2018, 17/01448, ECLI:NL:HR:2018:1316 (per la quale la pubblica amministrazione non può fare affidamento su sistemi algoritmici che non sono accessibili ai cittadini e alla pubblica amministrazione).

⁹³ Tribunale distrettuale dell'Aia, 2 marzo 2020, ECLI:NL:RBDHA:2020:865 (in cui si è concluso che la pubblica amministrazione può utilizzare un algoritmo di valutazione del rischio solo se esiste una chiara base giuridica per farlo); per un commento a questa decisione, si v. S. Elif Biber, *Between Humans and Machines*, cit., 201.

⁹⁴ Tribunale centrale del contenzioso amministrativo (*Juzgado Central de lo Contencioso Administrativo*), 31 dicembre 2021, n. 143, ECLI:ES:AN:2021:5863 (che ha confermato la legittimità dell'uso e della segretezza degli algoritmi utilizzati dalla pubblica amministrazione per scopi di pubblica utilità).

⁹⁵ Consiglio di Stato, 13 dicembre 2019, n. 8472; Consiglio di Stato, 8 aprile 2019, n. 2270 (secondo cui la pubblica amministrazione può avvalersi di algoritmi nella misura in cui questi ultimi siano trasparenti e impugnabili dinanzi ai tribunali). Su questa giurisprudenza, v. S. Elif Biber, *Between Humans and Machines*, cit., 203.

classificare le candidature delle persone in cerca di un'occupazione⁹⁶. Su richiesta della Corte costituzionale belga, la Corte di giustizia dell'Unione europea ha valutato l'irragionevolezza dell'uso da parte degli aeroporti europei di mezzi automatizzati per il trattamento dei dati dei passeggeri in arrivo da paesi extra UE e il loro confronto con banche dati internazionali di criminali nel contesto delle misure antiterrorismo, considerando come questi programmi davano luogo a cinque falsi positivi ogni sei segnalazioni⁹⁷.

Per ora, nulla di simile si è verificato nei paesi dell'Europa centrale e orientale. In Polonia, il Tribunale costituzionale ha deciso nel 2018 che le disposizioni lavoristiche sulla profilazione automatizzata dei disoccupati da parte degli uffici di collocamento erano incostituzionali, poiché non era possibile contestare i risultati di tale profilazione⁹⁸. L'uso di mezzi automatizzati per valutare il rispetto delle norme riguardo alla circolazione veicolare ha dato luogo a contenziosi limitati in Bulgaria⁹⁹ e a una denuncia di incostituzionalità in Serbia¹⁰⁰. A parte questi casi, i relatori provenienti da paesi dell'Europa centro-orientale non hanno segnalato altro contenzioso contro gli enti pubblici che si avvalgono dell'ADM. Tuttavia, data l'importanza e la centralità dell'uso dell'ADM da parte delle amministrazioni pubbliche, il numero sempre crescente di settori in cui può essere utilizzato e le sue implicazioni in materia di diritti umani, sembra ragionevole attendersi che anche in questi paesi aumenti progressivamente il numero delle liti riguardanti decisioni automatizzate.

7. Conclusioni

Pur essendo preliminari e ancora frammentari, i risultati fin qui raccolti consentono di individuare alcune traiettorie di convergenza e divergenza fra le giurisdizioni esaminate.

Il primo elemento di convergenza concerne la constatazione dalla quale abbiamo preso le mosse: lo stato automatizzato è in aumento ovunque. L'uso dell'ADM da parte delle amministrazioni pubbliche, comprese le

⁹⁶ Corte Suprema Amministrativa (*Verwaltungsgerichtshof*), 21 dicembre 2023, Ro 2021/04/0010-11 (in cui si sostiene che l'algoritmo non era conforme all'art. 22 del GDPR). Su questa decisione, si v. A. Ferrari Zumbini – P. Monaco, *Automated Decisions in Public Administration Processes: The Case of the Austrian Public Employment Service Algorithm*, in *Italian Journal of Public Law*, 17, 2025, 347 ss. La successiva decisione di merito ha però concluso che l'intervento umano fosse sufficiente a escludere l'applicabilità dell'art. 22 GDPR: Bundesverwaltungsgericht, 1° settembre 2025, W256 2235360-1/36E.

⁹⁷ CGUE, C-817/19, *Ligue des droits humains ASBL v. Conseil des ministres* (2022) (che ha confermato la legittimità dei controlli automatizzati nel loro complesso, pur criticando la mancanza di una revisione umana individuale dei risultati automatizzati).

⁹⁸ Corte costituzionale (*Trybunał Konstytucyjny*), 6 giugno 2018, fascicolo n. K 53/16.

⁹⁹ D. Marcheşa – S. Tsoneva, *The Law of the Algorithmic State in Bulgaria*, in *Italian Journal of Public Law*, 17, 2025, 529 ss. (con riferimento alle poche contestazioni presentate contro le multe automatiche emesse da dispositivi fissi).

¹⁰⁰ M. Milenković – S. Vračar – M. Pešterić, *The Law of the Algorithmic State in Serbia*, cit. (in cui si discute della legalità dell'uso del cosiddetto programma *Hawk Eye* da parte della polizia nella città di Belgrado, attualmente all'esame della Corte costituzionale serba).

autorità centrali e locali, le agenzie e le autorità indipendenti, è in crescita a tutte le latitudini e longitudini. Indipendentemente dal livello di coinvolgimento dello stato nell'industria della tecnologia digitale, le amministrazioni pubbliche sono ovunque attori di primo piano in questo campo.

Esiste anche una sostanziale convergenza per quanto riguarda i settori in cui si verifica questo processo. Quasi dappertutto, il passaggio allo stato digitale ha portato alla creazione del cosiddetto e-government, con un'identità digitale unica per i cittadini e portali online semi-centralizzati attraverso i quali le persone possono accedere direttamente a dati e documenti, ottenere certificati e licenze e gestire procedure di registrazione, iscrizione e simili. L'ADM è frequente soprattutto nei settori in cui i funzionari pubblici hanno poca o nessuna discrezionalità, come il pagamento delle tasse, l'erogazione delle prestazioni sociali, l'automazione dei servizi di trasporto e l'emissione di multe per determinate infrazioni alle regole sul traffico veicolare. Meno diffuse, ma comunque degne di nota, sono le sperimentazioni con tecnologie digitali e l'IA nei settori della sicurezza interna e della gestione delle frontiere, dell'istruzione, dell'agricoltura, del lavoro e della sanità. La convergenza diventa sempre meno evidente man mano che ci si allontana da questi settori fondamentali¹⁰¹.

Anche le motivazioni alla base di queste trasformazioni sembrano essere molto simili. Quasi ovunque si prevede che l'adozione delle tecnologie digitali migliorerà la qualità, l'efficienza e l'efficacia della pubblica amministrazione, oltre ad aumentare la trasparenza, la responsabilità, l'accessibilità e la parità di accesso ai servizi pubblici, a promuovere il coordinamento e la coerenza interna, a ridurre la burocrazia e i costi legati ai ritardi, a limitare la corruzione e, in ultima analisi, a migliorare la governance del sistema nella sua interezza. Allo stesso tempo, vi è una diffusa consapevolezza dei rischi significativi che l'automazione comporta. Molti dei relatori nazionali sono preoccupati per le potenziali implicazioni dell'ascesa dello stato intelligente in termini di sicurezza informatica, privacy, protezione dei dati e trasparenza nelle decisioni e nelle azioni pubbliche. In alcuni paesi, vi è un notevole timore che i progressi tecnologici possano causare disoccupazione, emarginare alcuni gruppi dalla vita pubblica e dai servizi pubblici e aumentare la dipendenza nazionale dai fornitori di tecnologia stranieri¹⁰².

Infine, in tutti gli ordinamenti giuridici esaminati, non vi è dubbio che i requisiti procedurali generalmente applicabili alle decisioni amministrative debbano essere adeguati e forse rafforzati quando le decisioni amministrative sono prese in modo automatizzato. L'idea che le decisioni amministrative algoritmiche debbano rispettare gli obblighi di protezione dei dati, la trasparenza, il diritto di ricevere spiegazioni, il dovere di supervisione umana e il diritto di ricorso sembra essere particolarmente radicata negli stati membri della UE, anche grazie al GDPR e all'AI Act. L'adozione di quest'ultimo suggerisce che presto ci sarà un grado ancora maggiore di uniformità tra i paesi dell'Unione, che potrebbe forse influenzare anche gli

¹⁰¹ Si v. sopra, par. 4.

¹⁰² Si v. sopra, par. 5.

stati confinanti grazie all'effetto Bruxelles¹⁰³.

Accanto alle tendenze di convergenza sopra descritte, esistono però anche molte divergenze, specie per quanto riguarda gli aspetti normativi, procedurali e istituzionali.

È evidente che i paesi analizzati si trovano in fasi diverse di implementazione e regolamentazione dell'automazione algoritmica e dell'IA nella pubblica amministrazione. Esiste un ampio divario tra paesi come gli Stati Uniti e l'Estonia, che dimostrano alti livelli di maturità digitale e servizi pubblici digitali avanzati, e stati, soprattutto nell'Europa sud-orientale, che stanno ancora costruendo le loro infrastrutture digitali e sono più lenti nell'adozione di sistemi di IA¹⁰⁴.

Non sorprende che queste differenze sembrino influenzare direttamente il livello di contenzioso contro lo stato automatizzato esistente in ciascuna giurisdizione. Pare che ci sia una correlazione diretta tra il grado di digitalizzazione dello stato e il numero di ricorsi presentati in tribunale contro l'amministrazione digitale, anche se questa generalizzazione va considerata con una certa cautela¹⁰⁵. Al contrario, il livello di digitalizzazione non pare correlato né alla portata delle riforme legislative riguardanti lo stato automatizzato né al numero di requisiti normativi che si applicano specificamente all'ADM amministrativa.

Per quanto concerne le riforme legislative, alcuni paesi hanno modificato le loro leggi generali sul procedimento amministrativo per consentire l'ADM in materia amministrativa o hanno adottato norme generali sull'e-government e la digitalizzazione della pubblica amministrazione. Tuttavia, molti altri paesi si sono astenuti dall'intervenire, lasciando che le sperimentazioni nell'utilizzo dell'ADM fossero regolate dalle norme generali esistenti¹⁰⁶. Sul fronte dei requisiti procedurali, si è riscontrato che, indipendentemente dal grado di sviluppo tecnologico e dagli approcci alla regolamentazione dell'ADM amministrativa, il numero e la severità dei requisiti per l'adozione dell'ADM sono più elevati negli stati membri della UE rispetto ai paesi non appartenenti alla UE, in gran parte a causa dell'esistenza del GDPR e dell'AI Act¹⁰⁷. Tuttavia, come illustra l'indagine sulla giurisprudenza realizzata nel par. VI, l'impatto di questa legislazione non deve essere sopravvalutato¹⁰⁸. I paesi della UE, a dispetto del comune quadro giuridico di riferimento, seguono modelli significativamente diversi se si tiene conto della giurisprudenza nazionale: mentre l'ADM pubblica è stata ripetutamente contestata nei tribunali dell'Europa occidentale, i tassi di contenzioso sono vicini allo zero nell'Europa centrale e orientale¹⁰⁹.

Se questa conclusione getta qualche ombra sull'effettivo livello di armo-

¹⁰³ Come già ricordato sopra, un altro potenziale strumento di armonizzazione in Europa è ora rappresentato dalla Convenzione quadro sull'intelligenza artificiale e i diritti umani, la democrazia e lo Stato di diritto, adottata nel 2024 dal Consiglio d'Europa.

¹⁰⁴ Stanford University, *The AI Index Report 2025*, cit.

¹⁰⁵ Si v. sopra, par. 6.

¹⁰⁶ Si v. sopra, par. 3.

¹⁰⁷ Si v. sopra, par. 5.

¹⁰⁸ Si v. sopra, par. 6.

¹⁰⁹ Si v. sopra, par. 6.

nizzazione giuridica in Europa, essa conferma anche la necessità di una ricerca comparatistica più approfondita riguardo al funzionamento pratico dei sistemi amministrativi nazionali.

Abstract

L'articolo presenta i primi risultati di un progetto di ricerca riguardo le decisioni amministrative automatizzate condotto in venti ordinamenti giuridici europei e negli Stati Uniti. L'articolo mira a far luce sul coinvolgimento dello stato non solo come regolatore, ma anche come utente e, occasionalmente, sviluppatore di nuove tecnologie. Esso perciò esplora il quadro giuridico applicabile alle decisioni amministrative automatizzate nelle giurisdizioni selezionate, i settori coinvolti, le condizioni alle quali tali decisioni possono essere assunte e le controversie derivanti dal loro utilizzo. Come emerge dall'articolo, esistono differenze sostanziali tra i sistemi giuridici in termini sia di approcci normativi che di tassi concreti di ricorso alle decisioni amministrative automatizzate. La conclusione principale del presente scritto è che queste differenze non sembrano essere correlate fra loro: gli approcci normativi prescelti sembrano avere un impatto limitato sulla misura effettiva di utilizzo di decisioni amministrative automatizzate, e viceversa.

This paper presents the results of a comparative research project on automated decision-making (ADM) carried out on twenty European legal systems and the United States. The paper aims to shed light on the involvement of the state not only as a regulator, but also as a major user and, occasionally, a developer of new technologies. It therefore explores the legal framework applicable to automated administrative decisions in the selected jurisdictions, the sectors involved, the conditions under which automated administrative decisions can be made, and the disputes arising from their use. As the paper shows, there are substantial differences between legal systems in terms of regulatory choices and of rates of adoption of administrative ADM. The main finding of the paper is that these differences do not appear to be correlated, in that regulatory choices have little impact on the public adoption of ADM, and vice versa.

Keywords

diritto amministrativo comparato - diritto e tecnologia - decisioni amministrative automatizzate - intelligenza artificiale - e-government

AI e amministrazione federale statunitense: strategie, cambi di rotta e implicazioni globali*

Laura Fabiano

Sommario

1. Intelligenza artificiale e politica federale: due visioni statunitensi a confronto. - 2 - La “cauta ambizione” della *policy* statunitense in tema di AI durante l'amministrazione democratica Biden tra *leadership* tecnologica e diritti fondamentali. - 3. DEI, MEI e intelligenza artificiale: il ripensamento della regolazione federale nel secondo mandato Trump. - 4. *Stargate Project*: la supremazia statunitense nell'intelligenza artificiale attraverso gli investimenti infrastrutturali. - 5. AI e divergenza regolatoria: l'effetto della seconda amministrazione Trump sull'equilibrio globale.

1. Intelligenza artificiale e politica federale: due visioni statunitensi a confronto

Tanto negli Stati Uniti quanto nel contesto europeo, la disciplina concernente l'ecosistema digitale si è caratterizzata storicamente per una prima lunga fase storica, connotata da tratti tendenzialmente liberistici, ove il tema maggiormente sentito era favorire la libera circolazione e la promozione dei servizi della società dell'informazione eliminando gli ostacoli allo sviluppo del commercio elettronico. Tale impianto seguiva al copioso dibattito (fra gli studiosi del tema) attinente alla considerazione dello spazio digitale quale (non)luogo anarchico o, al più, autogestito¹.

*L'articolo è stato sottoposto, in conformità al regolamento della Rivista, a referaggio “a doppio cieco”.

Il presente contributo è realizzato nell'ambito del PRIN 2020 “Decision-Making in the Age of Emergencies. New Paradigms in Recognition and Protection of Rights [DeMa]” – CUP C61B22002230001

¹ La dottrina giuridica si è interrogata, inizialmente, e per lungo tempo, sulla stessa praticabilità dell'estensione al mondo telematico del diritto territoriale vigente nel mondo reale e, dunque, sul c.d. “statuto costituzionale di Internet”. Tale questione dottrinale ha subito un'evoluzione storica, oramai trentennale, nel corso della quale da un primo approccio improntato all'idea della completa autonomia di Internet rispetto al mondo reale, si è approdati -grazie probabilmente al contributo rilevante della giurisprudenza di alcune Corti-, alla concezione piena della vigenza del diritto territoriale anche nel contesto telematico. Fra gli studiosi patrocinanti della concezione autonoma e anarchica del mondo digitale si richiama innanzitutto John Perry Barlow il quale con la sua *Declaration of Independence of the Cyberspace* del 1996 (in *eff.org*, 8 febbraio 1996) incarnava, sotto un profilo dottrinale, quelle posizioni ultraliberarie che si sono imposte nei primi anni di affermazione della rete *web* per le quali lo spazio digitale era, e doveva rimanere ad ogni costo, un (non) luogo di assoluta sfrenata libertà

La progressiva evidenza delle peculiarità tecniche connesse al contesto informatico ha tuttavia spinto a rivalutare la necessità di normare la dimensione tecnica del mezzo digitale e, dunque, a considerare l'esigenza di disciplinare con maggior rigore le questioni attinenti alla responsabilità del *provider* e l'attività algoritmica.

Con riguardo più strettamente al tema della normazione inerente all'intelligenza artificiale (AI) l'approccio Europeo e statunitense si è invece rivelato, sin da subito, distante dal liberismo classico e tuttavia reciprocamente divergente ove, rispetto all'importanza attribuita in Europa alla tutela dei diritti fondamentali individuali e collettivi, l'orientamento statunitense alla regolamentazione dell'AI si è caratterizzato per una maggiore attenzione verso la promozione della *leadership* economica nazionale².

Nondimeno, le politiche adottate negli Stati Uniti sono state profondamente influenzate dai cambiamenti amministrativi federali che si sono susseguiti nel tempo, in relazione ai quali si evidenzia un approccio repubblicano concentrato sulla supremazia tecnologica e sulla deregolamentazione ed un'impostazione democratica volta maggiormente a ricercare un equilibrio tra innovazione e protezione dei diritti individuali³.

individuale. Fra i sostenitori della *self regulation* del mondo digitale, «*lex specialis* convenzionale pensata per internet e partorita da internet» (M. Bassini, *Internet e libertà di espressione*, Prospettive Costituzionali e Sovranazionali, Roma, 2019, 46) figurano David Johnson e David Post i quali, in diversi scritti, (il più noto è *Law and Borders. The Rise of Law in Cyberspace*, in *Stanford Law Review*, 48, 5, 1996, 1367 ss.) contestano la possibilità di regolazione statale per il mondo digitale e sostengono l'autosufficienza regolatoria delle comunità digitali.

Parallelamente a tale tipologia di dibattito, la dottrina ha sviluppato un secondo filone di dissertazione concernente la sussistenza o meno di una dignità scientifica da riconoscersi allo studio delle dinamiche di *governance* e delle relazioni che si sviluppano nella rete *web*. Ciò a partire dalla provocazione di Frank Easterbrook (allora giudice della Corte di Appello degli Stati Uniti per il Settimo Circuito) il quale, nel 1996, in una conferenza sul tema *Law of Cyberspace* organizzata dall'Università di Chicago, contestò l'idea che il diritto digitale possedesse una sua intrinseca autonomia sostenendo che i suoi fenomeni dovevano essere studiati e analizzati attraverso il prisma delle regole generali. Sul fronte opposto si schierò con decisione Lawrence Lessing al quale, dunque, fa capo quel filone di pensiero che, riconoscendo alla dimensione digitale delle specificità irriducibili, riteneva che le caratteristiche del *web* imponessero una seria valutazione circa la possibilità delle categorie giuridiche tradizionali di permearne i confini e quindi di soddisfare pienamente le esigenze di regolazione e tutela che anche da tale (non)luogo promanano. Sul punto cfr. S. Rodotà, *Tecnopolitica. La democrazia e le nuove tecnologie della comunicazione*, Roma, 1997; V. Zeno-Zencovich, *Informatica ed evoluzione del diritto*, in *Diritto dell'informazione e dell'informatica*, 1, 2003; M. Bassini, *Internet e libertà di espressione*, cit; P. Costanzo, *Il fattore tecnologico e le sue conseguenze*, in *Rassegna parlamentare*, 4, 2012, 811 ss.; T.E. Frosini, *Tecnologie e libertà costituzionali*, in *Diritto dell'informazione e dell'informatica*, 3, 2003, 487 ss.; T.E. Frosini - O. Pollicino - E. Apa - M. Bassini (a cura di), *Diritti e libertà in Internet*, Firenze, 2017; M. Olivetti, *Diritti fondamentali e nuove tecnologie. Una mappa del dibattito italiano*, in *Revista Estudos Institucionais*, 6, 2, 2020, 395 ss. Sul tema dell'impatto dell'evoluzione tecnologica informatica sul diritto pubblico e sulla tenuta delle garanzie costituzionali si rinvia, inoltre, in generale, ai numerosi interessanti contributi presenti nel vol. 1 del *Liber Amicorum per Pasquale Costanzo* dedicato a *Costituzionalismo, Reti e Intelligenza Artificiale*, in *ConsultaOnline*, 2020.

² Cfr. sul punto L. Fabiano, *Il liberal-protezionismo digitale statunitense fra difesa della leadership nel mercato tecnologico e sicurezza nazionale*, in *DPCE online*, 60, 3, 2003. Cfr. anche L. Fabiano, *Le potenzialità manipolative della democrazia digitale fra interessi pubblici e poteri privati*, in *Rivista dell'informazione e dell'informatica*, 4/5, 2023, 597 ss.

³ Sul tema V. Lubello, *From Biden to Trump: Divergent and Convergent Policies in The Artificial Intelligence (AI) Summer*, in *DPCE online*, 1, 2025.

Ogni diversa amministrazione ha dunque promosso un certo livello di coinvolgimento del governo federale nella regolazione e promozione del settore; ciò che ha segnato la differenza è stata la finalità dell'intervento governativo. L'amministrazione Trump sembra infatti concepire l'intelligenza artificiale come uno strumento di proiezione della potenza nazionale e, dunque, le scelte politiche (e normative) che hanno caratterizzato (e che connotano tuttora) tale amministrazione nel detto ambito sembrano subordinate all'obiettivo della egemonia tecnologica americana; ciò anche a costo di ridurre gli spazi di regolazione e di neutralizzare i controlli etici e sociali. Diversamente, l'amministrazione Biden ha adottato un'impostazione più orientata alla *governance* responsabile dell'innovazione, cercando di bilanciare lo sviluppo tecnologico con la tutela dei diritti individuali e collettivi, come la *privacy*, la non discriminazione e la trasparenza. Nella detta prospettiva, dunque, la regolazione non è vista come un ostacolo alla competitività, bensì come una condizione necessaria affinché l'intelligenza artificiale possa essere compatibile con i valori democratici.

Nel confronto fra le scelte delle due amministrazioni governative si delinea pertanto un contrasto netto tra due visioni dell'intervento pubblico: una, ispirata alla logica della potenza; l'altra, alla logica della responsabilità democratica.

Tale grande divario, lampante a partire dalla prima amministrazione Trump (nell'adozione dell'*Executive Order* n. 13859 del 2019 intitolato emblematicamente "*Maintaining American leadership in AI*")⁴, si rende evidente, e trova conferma, nel paragone fra le scelte compiute dall'amministrazione democratica Biden con le più recenti decisioni assunte dal quarantasettesimo Presidente degli Stati Uniti d'America, Donald Trump, nel suo secondo mandato.

⁴ Il detto EO era finalizzato a promuovere investimenti federali nell'intelligenza artificiale come priorità per le agenzie governative, ad ottimizzare le risorse federali esistenti per la valorizzazione dello sviluppo dell'intelligenza artificiale americana e ad investire nella formazione di esperti e nella protezione delle tecnologie statunitensi. Esso è stato adottato a seguito del *White House Summit on AI for American Industry* del 2018. La strategia inaugurata dalla presidenza federale USA nel 2019 era collegata, almeno idealmente, all'ambizioso piano cinese del 2017, con il quale Pechino dichiarava l'intenzione di diventare *leader* mondiale nell'AI entro il 2030. La reazione statunitense si concretizzò in una politica volta a consolidare il primato tecnologico americano, dando priorità alla sicurezza nazionale e alla protezione del mercato interno piuttosto che alla tutela dei diritti individuali nell'uso dell'AI.

Parallelamente, il Congresso degli Stati Uniti confermò questa linea politica con l'emanazione della sezione 1051 del *John S. McCain National Defense Authorization Act* per l'anno fiscale 2019, istituendo la *National Security Commission on AI*. Tale organismo aveva il compito di promuovere lo sviluppo dell'intelligenza artificiale, dell'apprendimento automatico e delle tecnologie connesse per la sicurezza e la difesa nazionale. Successivamente, nel 2020, venne proposto il *National AI Initiative Act*, un atto normativo per pianificare e coordinare le attività federali a sostegno della ricerca nell'ambito dell'AI. Il Congresso sceglieva dunque di seguire l'indirizzo presidenziale mentre non venivano al contempo approvate altre proposte normative maggiormente orientate a valutare i rischi connessi allo sviluppo dell'AI ed a considerare l'eventuale quadro delle responsabilità ad essa connesse.

2. La “cauta ambizione” della *polycy*statunitense in tema di AI durante l'amministrazione democratica Biden tra *leadership* tecnologica e diritti fondamentali

Come anticipato, l'amministrazione Biden ha mostrato, sin da subito, una crescente attenzione nei confronti dell'impatto individuale e sociale dell'intelligenza artificiale, pur non rinunciando alla competitività tecnologica ed alla sicurezza nazionale.

Nel 2022, sulla spinta dell'amministrazione democratica, veniva infatti pubblicato il “*Blueprint for an AI Bill of Rights*”⁵, un documento dell'*Office of Science Technology Policy* della Casa Bianca⁶ che stabiliva delle linee guida per garantire un utilizzo dell'AI equo, sicuro e rispettoso dei diritti fondamentali. Sebbene non avesse valore legislativo, il documento, in sostanza, poneva cinque principi alla guida della progettazione e dell'uso responsabile dell'intelligenza artificiale (*Safe and Effective Systems, Algorithmic Discrimination Protections, Data Privacy, Notice and Explanation, Human Alternatives, Consideration and Fallback*)⁷, fornendo, a margine del documento, una definizione per i termini chiave utilizzati.

L'orientamento della Presidenza Biden in tema di AI si è successivamente sviluppato in una serie di provvedimenti finalizzati a consolidare questa nuova direzione della politica americana (una fase che è stata definita la “*to do list macro-phase*”)⁸; ciò in primo luogo attraverso l'adozione, nel 2023, dell'*Executive Order* n. 14110 intitolato “*Safe, secure and Trustworthy Development and Use of Artificial Intelligence*”. Anche in questo caso, gli obiettivi politici delineati riguardavano tanto la promozione della concorrenza e la garanzia della competitività globale degli Stati Uniti nel settore dell'intelligenza artificiale quanto la prevenzione delle minacce alle libertà civili e alla sicurezza nazionale rese possibili da tale tecnologia⁹.

⁵ Il documento è attualmente reperibile all'url [biodiritto.org](https://www.biodiritto.org).

⁶ L'*Office of Science Technology Policy* è un ufficio istituito con il *National Science and Technology Policy, Organization, and Priorities Act* del 1976 con l'obiettivo di supportare la Presidenza e, più in generale, l'Ufficio Esecutivo del Presidente con attività di consulenza sugli aspetti scientifici, ingegneristici e tecnologici dell'economia, della sicurezza nazionale, della salute, delle relazioni estere, dell'ambiente e del recupero ed utilizzo tecnologico delle risorse. Cfr. E.G. Belvins, *The Office of Science and Technology Policy (OSTP): Overview and Issues for Congress*, Crs Report for Congress, R 47410, 2023.

⁷ Sul documento cfr. E. Hine - L. Floridi, *The Blueprint for an AI Bill of Rights: In Search of Enaction, at Risk of Inaction*, in *Minds and Machines*, 2023.

⁸ Tale la definisce V. Lubello, *From Biden to Trump: Divergent and Convergent Policies in The Artificial Intelligence (AI) Summer*, cit.

⁹ Gli obiettivi politici delineati dall'EO includevano: promuovere la concorrenza e l'innovazione nel settore dell'intelligenza artificiale; tutelare i diritti civili e del lavoro e proteggere i consumatori e la loro *privacy* dai danni causati dall'intelligenza artificiale; specificare le politiche federali che regolano l'approvvigionamento e l'uso dell'intelligenza artificiale; sviluppare di sistemi *watermarking* (di filigrana) per i contenuti generati dall'intelligenza artificiale e prevenzione del furto di proprietà intellettuale derivante dall'uso di modelli generativi; mantenere la posizione della nazione come *leader* mondiale nell'intelligenza artificiale. Sul detto EO cfr. C. Sbailò, *Executive Order 14110 Governing Artificial Intelligence: Technological Leadership and Regulatory Challenges in an Era of Exponential*

Significativamente l'EO 14110 richiama, nel proprio testo, il *NIST AI Risk Management Framework (AI RMF)*¹⁰, un documento pubblicato nel gennaio 2023 dal *National Institute of Standards and Technology* (un'agenzia federale del Dipartimento del Commercio degli Stati Uniti) con l'obiettivo di fornire una guida operativa e flessibile per aiutare organizzazioni ed aziende a comprendere, valutare, gestire e mitigare i rischi associati all'uso dei sistemi di intelligenza artificiale.

Il documento si segnalava in primo luogo in relazione al suo processo d'adozione in quanto risulta essere il frutto di un processo partecipativo significativamente esteso che ha coinvolto Agenzie governative federali¹¹; industrie private grandi e piccole¹²; Accademia e centri di ricerca¹³; Organizzazioni della società civile ed ONG¹⁴. Nell'approvazione del documento, inoltre, il NIST ha aperto consultazioni pubbliche e richiesto contributi attraverso *Requests for Information (RFI)*, *public workshops* e *listening sessions*, consentendo a cittadini, sviluppatori indipendenti e altri *stakeholders* di intervenire attraverso: bozze pubbliche (*Drafts*); *Workshops* tematici; Analisi comparative di modelli internazionali (es. *OCSE AI Principles*, *UNESCO Recommendation on AI Ethics*, *AI Act* europeo).

Il *Framework* posto in essere attraverso tale percorso, che intendeva come suoi destinatari non solo gli sviluppatori, ma anche utilizzatori, regolatori, fornitori, imprese e istituzioni pubbliche, si articolava in quattro elementi fondamentali: *Govern, Map, Measure, Manage*¹⁵.

Growth, in *DPCE online*, sp. 3, 2024, 275 ss.

¹⁰ Per un commento alle bozze dell'*AI Risk Management Framework* cfr. F.M.R. Livelli, *NIST AI Risk Management Framework (RMF): come garantire l'affidabilità dell'IA secondo gli Usa*, in *Agenda Digitale*, 15 settembre 2022.

¹¹ Ad es. il Dipartimento del Commercio, il Dipartimento della Difesa, il *Department of Homeland Security*, agenzie sanitarie, enti di ricerca pubblica.

¹² Tra cui grandi aziende tecnologiche quali Google, Microsoft, IBM, Amazon, Meta, sia *start-up* del settore AI, associazioni di categoria (ad es. *U.S. Chamber of Commerce*, *Business Roundtable*), ma anche aziende appartenenti a settori non strettamente tech (ad es. bancario, manifatturiero, assicurativo).

¹³ Università statunitensi e internazionali (come Stanford, MIT, Harvard, Carnegie Mellon), *think tank* (Brookings, Center for Democracy & Technology, AI Now Institute), esperti di etica dell'AI e governance tecnologica.

¹⁴ Ad es. *L'Electronic Frontier Foundation (EFF)*, *l'American Civil Liberties Union (ACLU)*, *il Partnership on AI*, *Future of Life Institute*.

¹⁵ Le dette fasi non sono rigide o lineari, ma costituiscono un ciclo continuo e dinamico per gestire in modo responsabile i rischi legati all'utilizzo dell'intelligenza artificiale. La prima fase attiene al "governo del rischio" ed è orientata a creare una struttura organizzativa e culturale interna che permetta di affrontare in modo sistematico i rischi dell'AI (definire in anticipo chi è responsabile della gestione del rischio nell'organizzazione); assicurarsi che i valori etici (come equità, trasparenza, responsabilità siano integrati in tutte le fasi di sviluppo dell'AI); favorire il coinvolgimento degli *stakeholders* (comunità, utenti, esperti) nelle decisioni importanti; stabilire procedure e risorse per monitorare i rischi nel tempo. La seconda fase attiene alla "mappatura del contesto e dei rischi" al fine di comprendere il sistema AI nel suo ambiente d'uso e identificare i possibili rischi per persone, gruppi o società. La terza fase riguarda la "misurazione dei rischi" attraverso l'utilizzo di strumenti oggettivi e/o qualitativi (es. tassi di errore, *bias* statistici, performance predittiva). L'ultima fase riguarda la "Gestione e la mitigazione dei rischi" attraverso l'adozione di modifiche tecniche (es. migliorare il *dataset*, cambiare l'algoritmo) e salvaguardie umane (come la supervisione o il diritto all'intervento umano) e procedure di revisione continua.

A ridosso della fine del proprio mandato, la Presidenza Biden ha adottato ulteriori documenti significativi: il 24 ottobre del 2024 è stato adottato un *Memorandum* sulla sicurezza nazionale dedicato all'AI generativa volto a bilanciare, ancora una volta, il tema della sicurezza individuale e nazionale con gli obiettivi di sviluppo tecnologico¹⁶; il 14 gennaio del 2025, a sei giorni dal termine del proprio incarico, il Presidente Biden ha adottato infine l'*Executive Order* n. 14141 dal titolo "*Advancing United States Leadership in Artificial Intelligence Infrastructure*". In tale documento la Presidenza democratica si occupava di un altro aspetto importante della decisione politica concernente l'AI, ovvero del tema delle necessità infrastrutturali che riguardano tale tecnologia ed in particolare della necessità di costruire *data center* e di adattare la struttura energetica del Paese a tali sviluppi.

L'EO 14141 poneva alcuni principi guida al fine di bilanciare le esigenze di sviluppo infrastrutturale del Paese in tema di AI con gli altrettanto importanti obiettivi, fra cui la promozione dell'energia pulita, la garanzia di un ecosistema tecnologico competitivo e aperto alle grandi, ma anche alle piccole aziende e tutela dei consumatori e delle persone le cui comunità vivono nei pressi di tali realtà infrastrutturali.

L'ordine mirava a consolidare la *leadership* degli Stati Uniti nel campo dell'intelligenza artificiale attraverso lo sviluppo di infrastrutture nazionali avanzate, proteggendo la sicurezza nazionale e la competitività economica, e promuovendo l'uso di energia pulita.

Le principali direttive riguardavano: la sicurezza nazionale (favorendo il rafforzamento della capacità degli Stati Uniti di sviluppare sistemi AI senza dipendere da infrastrutture straniere e proteggendo i laboratori e le catene di fornitura); la competitività economica (promuovendo la garanzia di un ecosistema di AI aperto, innovativo e accessibile anche a piccole imprese e sviluppatori indipendenti); la sostenibilità ecologica (incentivando la costruzione di *data center* basati su energia rinnovabile -geotermica, nucleare, solare, eolica - per ridurre le emissioni); la sostenibilità economica (evitando aumenti dei costi energetici per i consumatori e facendo sì che le aziende di AI finanziassero le infrastrutture necessarie); il beneficio per i lavoratori e le comunità coinvolte (caldeggiando la garanzia di standard di lavoro elevati, di coinvolgimento delle comunità locali nello sviluppo delle infrastrutture e di mitigazione dei potenziali impatti negativi).

L'ordine forniva quindi un quadro strategico per assicurare che l'AI di frontiera¹⁷ venisse sviluppata negli Stati Uniti in modo sicuro, competitivo e sostenibile.

Si rende evidente, anche in quest'ultima azione presidenziale, l'attenzione dell'amministrazione democratica nei confronti della ricerca di un necessario bilanciamento fra innovazione tecnologica e tutela di tutti gli altri interessi in gioco nell'ambito dello sviluppo (in questo ultimo caso, infrastrutturale) dell'AI. In particolare, la preoccupazione nei confronti della tutela ambientale e dei diritti delle comunità coinvolte da tali investimenti strategici risultavano essere gli elementi fondamentali da considerare al

¹⁶ Sul detto memorandum cfr. R. Panetta, IA, *Ecco il piano Biden: l'approccio pratico degli USA lezione per l'UE*, in *Agenda Digitale*, 30 ottobre 2024.

¹⁷ Sul concetto di *Frontier Ai* cfr. C. Bullock et al., *Legal Considerations for Defining "Frontier Model"*, in *Institute for Law & A.I., Working Paper 2*, 2024, in *law-ai.org*.

fine di adottare decisioni su uno sviluppo dell'AI equilibrato e di lungo periodo.

L'approccio regolatorio che sembra aver caratterizzato l'amministrazione democratica Biden, pur caratterizzato da una tendenza piuttosto chiara volta agli obiettivi di tutela dei diritti citata è, in ogni caso molto misurata, in quanto orientata ad evitare di soffocare l'innovazione tecnologica in un settore percepito come strategico per la competitività globale ed a preservare la *leadership* americana nel settore rispetto a concorrenti sistemici come la Cina. Tale tendenza regolatoria si percepisce nel suo proporsi principalmente quale strategia regolatoria volta a coordinare gli attori federali, pubblici e privati, attorno a principi comuni, piuttosto che a vincoli giuridici stringenti e nel suo svolgersi perlopiù attraverso strumenti tipici della *soft law*, quali *Executive Orders*, *Memorandum*, Documenti guida (es. *Blueprint for an AI Bill of Rights* e *AI RMF*).

Tali strumenti non sono mirati a produrre effetti giuridicamente vincolanti per il settore privato, salvo in ambiti molto specifici (si pensi al contesto degli appalti federali o alla sicurezza nazionale) ed il loro principale scopo attiene ad effetti di "*moral suasion*" principalmente orientando il comportamento degli sviluppatori e delle imprese, promuovendo *best practices*, creando consenso attorno a *standard* etici e tecnici e ponendo le basi per una successiva regolazione più solida. Tale circostanza segna una profonda differenza rispetto all'approccio europeo il quale, ugualmente diretto a bilanciare l'innovazione tecnologica con la tutela dei diritti individuali e collettivi ha intrapreso un percorso normativo di *hard law* (seppur dopo una lunga esperienza e un marcato intreccio di normazione di *soft law* e di coregolazione) attraverso l'adozione di normative importanti quali il *Digital Services Act*¹⁸, il *Digital Market Act*¹⁹ e soprattutto l'*AI Act*²⁰.

3. DEI, MEI e intelligenza artificiale: il ripensamento della regolazione federale nel secondo mandato Trump

Una decisa inversione di rotta si registra, nondimeno, nel gennaio del 2025 a pochi giorni dall'insediamento della seconda amministrazione Trump. Il

¹⁸ Regolamento (UE) 2022/2065 del Parlamento Europeo e del Consiglio del 19 ottobre 2022 relativo a un mercato unico dei servizi digitali e che modifica la direttiva 2000/31/CE (regolamento sui servizi digitali).

¹⁹ Regolamento (UE) 2022/1925 del Parlamento Europeo e del Consiglio del 14 settembre 2022 relativo a mercati equi e contendibili nel settore digitale e che modifica le direttive (UE) 2019/1937 e (UE) 2020/1828 (regolamento sui mercati digitali). Su regolamenti detti cfr. F. Zorzi Giustiniani, *L'Unione Europea e la regolamentazione del digitale: il Digital Services Package e il codice di buone pratiche sulla disinformazione*, in *Nomos. Le Attualità nel Diritto*, 2, 2022.

²⁰ Regolamento (UE) 2024/1689 del Parlamento europeo e del Consiglio, del 13 giugno 2024, che stabilisce regole armonizzate sull'intelligenza artificiale e modifica i regolamenti (CE) n. 300/2008, (UE) n. 167/2013, (UE) n. 168/2013, (UE) 2018/858, (UE) 2018/1139 e (UE) 2019/2144 e le direttive 2014/90/UE, (UE) 2016/797 e (UE) 2020/1828 (regolamento sull'intelligenza artificiale). Sulla differenza fra l'impianto normativo statunitense e quello europeo si rinvia a L. Fabiano, *Le potenzialità manipolative della democrazia digitale fra interessi pubblici e poteri privati*, cit.

47° Presidente degli Stati Uniti d'America inaugura difatti il proprio secondo mandato con un numero sorprendente di *Executive Orders* (sugli argomenti più diversi e tuttavia, molti, significativamente, intrecciati fra loro)²¹. In tema di regolazione dell'AI la presidenza Trump adotta diversi EO e dà vita a significativi progetti che si pongono in netta antitesi con gli obiettivi e le strategie dell'amministrazione precedente.

Già il 23 gennaio 2025, il 47° Presidente emette l'EO n. 14179 "*Removing Barriers to American Leadership in Artificial Intelligence*" con il quale sostituisce il precedente n. 14110 adottato da Biden, (già annullato il 20 gennaio 2025 con Ordine Esecutivo n. 14148 "*Initial Rescissions of Harmful Executive Orders and Actions*").

L'Order di Trump si propone in netta controtendenza rispetto all'enfasi posta dall'amministrazione Biden sulla supervisione, sulla mitigazione del rischio e sull'equità, delineando un quadro incentrato sulla deregolamentazione e sulla promozione dell'innovazione dell'intelligenza artificiale come mezzo per mantenere il predominio globale degli Stati Uniti.

L'EO 14179 è difatti esplicitamente finalizzato a «consolidare la ... posizione [degli USA] di leader globale nell'IA» ed in esso il Presidente Trump pone come finalità essenziale, da perseguire nella politica sull'IA, «il dominio globale dell'intelligenza artificiale dell'America al fine di promuovere la prosperità umana, la competitività economica e la sicurezza nazionale» (sez. 2). In consonanza con tali obiettivi l'EO stabilisce che «l'Assistente del Presidente per la scienza e la tecnologia (APST), il consulente speciale per l'intelligenza artificiale e la crittografia e l'Assistente del Presidente per gli affari di sicurezza nazionale (APNSA) esamineranno identificheranno tutte le azioni intraprese ai sensi dell'Ordine esecutivo 14110 che sono o potrebbero essere incoerenti con, o presentano ostacoli alla, politica stabilita nella sezione 2 e che, per qualsiasi azione identificata, i responsabili delle agenzie, come appropriato e coerente con la legge applicabile, sospenderanno, rivedranno o annulleranno tali azioni, o proporranno di sospendere, rivedere o annullare tali azioni».

L'EO del 23 gennaio 2025 prevede inoltre che «Entro 180 giorni dal presente ordine, l'Assistente del Presidente per la scienza e la tecnologia (APST), il Consigliere speciale per l'intelligenza artificiale e la crittografia e l'Assistente del Presidente per gli affari di sicurezza nazionale (APNSA), in coordinamento con l'Assistente del Presidente per la politica economica, l'Assistente del Presidente per la politica interna, il Direttore dell'Ufficio di gestione e bilancio (Direttore dell'OMB) e i responsabili dei dipartimenti esecutivi e delle agenzie (agenzie) che l'APST e l'APNSA ritengono pertinenti, svilupperanno e presenteranno al Presidente un piano d'azione per realizzare la politica stabilita nella sezione 2 del presente ordine».

Un passaggio significativo nell'EO del 23 gennaio risiede certamente in un inciso iniziale ove, nell'elencare gli scopi del documento, si afferma che è necessario «sviluppare sistemi di IA che siano liberi da pregiudizi ideologici o agende sociali».

Tale espressione collega certamente il detto EO ad altre coeve azioni pre-

²¹ Sul tema cfr. A. Baraggia - D. Camoni, *Gli "Executive Orders" nella seconda Presidenza Trump: verso un nuovo equilibrio nella separazione dei poteri?*, in *Diritto Pubblico Comparato ed Europeo*, 2, 2025, 241 ss.

sidenziali attinenti maggiormente al tema dell'antidiscriminazione.

Come noto, dal giorno del proprio insediamento alla Casa Bianca, il presidente Trump ha infatti adottato numerose azioni ed ordini esecutivi volti a realizzare un deciso passaggio dal paradigma DEI (*Diversity, Equality, Inclusion*) al paradigma MAI (*Merit, Excellence, Intelligence*).

Emblematico a tal proposito è l'EO n. 14151 adottato il 20 gennaio 2025 "*Ending Radical And Wasteful Government DEI Programs And Preferencing*" cui viene affiancato, il medesimo giorno l'EO "*Initial Rescissions of Harmful Executive Orders and Ations*". A seguito di questi due provvedimenti, il direttore *ad interim* dell' *Office of Personnel Management* (OPM) ha adottato delle linee guida di attuazione alle agenzie per questi ordini esecutivi in base alle quali è stato ordinato a tutte le amministrazioni federali di mettere i dipendenti degli uffici DEIA (uffici e sotto-unità di ogni agenzia funzionalizzate alla predisposizione di iniziative e programmi di Diversità, Equità, Inclusione e Accessibilità) in congedo amministrativo entro le 17:00 del 22 gennaio, nell'attesa di provvedere all'adozione delle misure idonee a porre fine a tutti gli uffici e programmi DEIA. Le agenzie sono state inoltre incaricate di inviare un avviso a tutti i dipendenti per informarli della chiusura di questi uffici e chiedere loro se fossero a conoscenza di «tentativi di mascherare questi programmi utilizzando un linguaggio codificato o impreciso». Il promemoria dell'OPM ha ordinato inoltre alle agenzie di rimuovere tutto il materiale correlato a DEIA dai loro siti Web e annullare qualsiasi formazione o contratto correlato a programmi e uffici DEIA. Le agenzie sono state inoltre incaricate di inviare all'OPM un elenco di tutti i dipendenti negli uffici DEIA. Entro il 31 gennaio, le agenzie dovevano inviare all'OPM un «piano scritto per l'esecuzione di un'azione di riduzione del personale riguardante i dipendenti che lavorano in un ufficio DEIA». L'EO ordina inoltre alle agenzie di segnalare all'OPM informazioni su eventuali beneficiari federali che si impegnano in attività correlate a DEI, DEIA o "giustizia ambientale".

Il 21 gennaio 2025 è stata inoltre adottata l'azione presidenziale "*Ending Illegal Discrimination and restoring merit based opportunity*" con la quale il Presidente ha ordinato «a tutti i dipartimenti esecutivi ed alle agenzie di porre fine a tutte le preferenze, mandati, politiche, programmi, attività, linee guida, regolamenti, azioni di esecuzione, ordini di consenso e requisiti discriminatori e illegali», interpretando come discriminatoria qualunque azione finalizzata all'inclusività ed al perseguimento dell'eguaglianza sostanziale²².

²² Con il detto OE il Presidente Trump ha revocato numerosissimi OE ed azioni presidenziali del suo predecessore fra cui: (i) l'OE 12898 dell'11 febbraio 1994 (Azioni federali per affrontare la giustizia ambientale nelle popolazioni minoritarie e nelle popolazioni a basso reddito);

(ii) l'OE 13583 del 18 agosto 2011 (Istituzione di un'iniziativa coordinata a livello governativo per promuovere la diversità e l'inclusione nella forza lavoro federale);

(iii) l'OE 13672 del 21 luglio 2014 (Ulteriori emendamenti all'Ordine esecutivo 11478, Pari opportunità di lavoro nel governo federale, e all'OE 11246, Pari opportunità di lavoro); e

(iv) il Memorandum presidenziale del 5 ottobre 2016 (Promozione della diversità e dell'inclusione nella forza lavoro della sicurezza nazionale).

L'EO ha inoltre sancito che:

(b) Il processo di appalto federale deve essere semplificato per migliorare la velocità e l'efficienza, ridurre i costi e richiedere ai contraenti e ai subappaltatori federali di rispettare

L'OE incentiva inoltre il settore privato "a porre fine alla discriminazione e alle preferenze illegali in materia di DEI"²³.

le leggi statunitensi sui diritti civili. Di conseguenza:

(i) L'OE 11246 del 24 settembre 1965 (Pari opportunità di lavoro) è revocato. Per 90 giorni dalla data di questo ordine, i contraenti federali possono continuare a conformarsi allo schema normativo in vigore il 20 gennaio 2025.

(ii) L'*Office of Federal Contract Compliance Programs* all'interno del *Department of Labor* cesserà immediatamente di:

(A) Promuovere la "diversità";

(B) Ritenere i contraenti e i subappaltatori federali responsabili dell'adozione di "azioni affermative"; e

(C) Consentire o incoraggiare i contraenti e i subappaltatori federali a impegnarsi nell'equilibrio della forza lavoro in base a razza, colore, sesso, preferenza sessuale, religione o origine nazionale.

(iii) In conformità con l'OE 13279 del 12 dicembre 2002 (Pari protezione delle leggi per le organizzazioni religiose e comunitarie), le pratiche di impiego, approvvigionamento e appalto dei contraenti e subappaltatori federali non devono considerare razza, colore, sesso, orientamento sessuale, religione o origine nazionale in modi che violino le leggi sui diritti civili della nazione.

(iv) Il responsabile di ciascuna agenzia deve includere in ogni contratto o concessione di sovvenzione:

(A) Un termine che richiede alla controparte contrattuale o al beneficiario della sovvenzione di accettare che la sua conformità in tutti gli aspetti a tutte le leggi federali antidiscriminazione applicabili è sostanziale per le decisioni di pagamento del governo ai fini della sezione 3729(b)(4) del titolo 31, Codice degli Stati Uniti; e

(B) Un termine che richiede a tale controparte o beneficiario di certificare di non gestire alcun programma che promuova la DEI che violi alcuna legge federale antidiscriminazione applicabile.

(c) Il Direttore dell'Ufficio di gestione e bilancio (OMB), con l'assistenza del Procuratore generale come richiesto, dovrà:

(i) Rivedere e modificare, se del caso, tutti i processi, le direttive e le linee guida dell'intero governo;

(ii) Eliminare i riferimenti ai principi DEI e DEIA, sotto qualsiasi nome possano apparire, dalle procedure federali di acquisizione, appalti, sovvenzioni e assistenza finanziaria per semplificare tali procedure, migliorare la velocità e l'efficienza, ridurre i costi e rispettare le leggi sui diritti civili; e

(iii) porre fine a tutti i mandati, requisiti, programmi o attività simili su "diversità", "equità", "processo decisionale equo", "distribuzione equa di assistenza finanziaria e tecnica", "promozione dell'equità" e simili, a seconda dei casi.

²³ A tal fine (a) I responsabili di tutte le agenzie, con l'assistenza del Procuratore generale, adotteranno tutte le misure appropriate in merito alle operazioni delle loro agenzie per promuovere nel settore privato la politica di iniziativa individuale, eccellenza e duro lavoro identificata nella sezione 2 dell'Order. (b) il Procuratore generale, entro 120 giorni dell'Order, in consultazione con i responsabili delle agenzie competenti e in coordinamento con il Direttore dell'OMB, invierà un rapporto all'Assistente del Presidente per la politica interna contenente raccomandazioni per l'applicazione delle leggi federali sui diritti civili e l'adozione di altre misure appropriate per incoraggiare il settore privato a porre fine alla discriminazione e alle preferenze illegali, inclusa la DEI. La relazione deve contenere una proposta di piano strategico di applicazione che identifichi:

(i) i settori chiave di interesse all'interno della giurisdizione di ciascuna agenzia;

(ii) i professionisti DEI più eclatanti e discriminatori in ciascun settore di interesse;

(iii) Un piano di misure o passaggi specifici per scoraggiare i programmi o i principi DEI (siano essi specificamente denominati "DEI" o in altro modo) che costituiscono discriminazioni o preferenze illegali. Come parte di questo piano, ogni agenzia identificherà fino a nove potenziali indagini di conformità civile di società quotate in borsa, grandi società o associazioni senza scopo di lucro, fondazioni con attività di 500

L'EO si occupa infine di smantellare l'impianto DEI in ambito scolastico con particolare riguardo all'istruzione superiore prevedendo, nella sez. 5, che «Entro 120 giorni dal presente ordine, il Procuratore generale e il Segretario dell'istruzione emetteranno congiuntamente una guida a tutte le agenzie educative statali e locali che ricevono fondi federali, nonché a tutte le istituzioni di istruzione superiore che ricevono sovvenzioni federali o partecipano al programma federale di assistenza per prestiti agli studenti ai sensi del Titolo IV dell'*Higher Education Act*, 20 USC 1070 e seguenti, in merito alle misure e alle pratiche richieste per conformarsi alla decisione della Corte suprema *Students for Fair Admissions, Inc. v President and Fellows of Harvard College*»²⁴.

È evidente come il tema del passaggio dal paradigma DEI al paradigma MEI si intrecci strettamente con gli obiettivi di sviluppo di «sistemi di IA che siano liberi da pregiudizi ideologici o agende sociali». Con tale espressione l'EO di Trump sembra sottendere infatti ad un modo di guardare al tema della discriminazione algoritmica del tutto diverso rispetto agli approcci che hanno caratterizzato la regolamentazione dell'IA fino ad ora. Sul tema, nel contesto federale statunitense, nel 2019 (H.R. 2231), nel 2022 (S. 3572) e poi nuovamente 2023 (H.R. 5628) è stata avanzata una proposta normativa al Congresso denominata *Algorithmic Accountability Act*²⁵. Obiettivo del testo era fronteggiare i rischi relativi a possibili discriminazioni e violazioni della *privacy* in relazione all'utilizzo di sistemi di intelligenza artificiale, richiedendo la redazione di analisi di impatto alle imprese che utilizzassero determinati sistemi, prevedendo obblighi di rendicontazione alla *Federal Trade Commission* (FTC) e stabilendo l'elaborazione da parte della FTC di regolamenti di implementazione e linee guida.

I più recenti *Executive Orders* di Trump si pongono pertanto in esplicito contrasto con il percorso finora intrapreso.

L'approccio della nuova amministrazione federale è confermato dalle parole del vicepresidente federale Usa J.D. Vance il quale, al vertice sull'intelligenza artificiale organizzato a Parigi nel febbraio del 2025, ha affermato chiaramente che una “regolamentazione eccessiva” potrebbe paralizzare il settore dell'intelligenza artificiale in rapida crescita²⁶. In relazione a tale

milioni di dollari o più, associazioni mediche e forensi statali e locali e istituzioni di istruzione superiore con dotazioni superiori a 1 miliardo di dollari;

(iv) Altre strategie per incoraggiare il settore privato a porre fine alle discriminazioni e alle preferenze DEI illegali e a conformarsi a tutte le leggi federali sui diritti civili;

(v) Contenzioso che sarebbe potenzialmente appropriato per cause legali federali, interventi o dichiarazioni di interesse; e (vi) Potenziale azione regolamentare e guida sub-regolamentare.

²⁴ *Students for Fair Admissions, Inc. v President and Fellows of Harvard College*, 600 US 181 (2023); in detta sentenza, come noto, la Corte suprema degli Stati Uniti ha sostanzialmente dichiarato l'incostituzionalità dell'uso, da parte delle università, delle *affirmative actions* nel processo di selezione degli studenti. Per un commento alla decisione si rinvia a D.E. Bernstein, *Students for Fair Admissions and the End of Racial Classification as We Know It*, in *Cato Supreme Court Review*, 2022/2023.

²⁵ B.C. Cheong, *Transparency and accountability in AI systems: safeguarding wellbeing in the age of algorithmic decision-making*, in *Frontiers in Human Dynamics*, 6, 3 July 2024.

²⁶ F. Madison, *JD Vance al Summit sull'Intelligenza Artificiale di Parigi: un monito contro l'iper-regolamentazione*, in *Rivista Ai*, 11 febbraio 2025.

impostazione non sorprende che gli Stati Uniti (insieme con la Gran Bretagna) non abbiano sottoscritto gli accordi raggiunti nel detto terzo vertice sull'intelligenza artificiale, sottoscritti invece da Europa e Cina.

L'EO di Trump diverge dal precedente di Biden anche con riguardo al tema dello sviluppo della forza lavoro e della formazione. L'EO di Biden, infatti, dedicava delle risorse all'attrazione e alla formazione di talenti AI, ampliando i percorsi di visto per lavoratori qualificati e promuovendo partnership pubblico-private per la ricerca e lo sviluppo AI. L'EO di Trump, invece, non include disposizioni specifiche relative alla forza lavoro e sembra presumere che la riduzione della supervisione federale consentirà naturalmente l'innovazione e la crescita dei talenti nel settore privato.

In relazione alle differenze rintracciabili fra l'approccio di Biden e quello di Trump nella disciplina dell'AI si è inoltre posto un significativo problema di tutela della sicurezza nazionale. Ciò considerando che mentre l'EO di Biden imponeva un'ampia cooperazione interagenzia per valutare i rischi che l'IA pone ai sistemi critici per la sicurezza nazionale, alle infrastrutture informatiche e alla biosicurezza (e richiedeva pertanto ad alcune agenzie, come il Dipartimento dell'Energia e il Dipartimento della Difesa, di condurre valutazioni dettagliate delle potenziali minacce dell'IA, incluso l'uso improprio dell'IA per lo sviluppo di armi chimiche e biologiche), la politica prescelta da Trump è invece connotata da una spiccata semplificazione della governance e dalla riduzione della supervisione federale nello sviluppo dell'AI.

Infine, l'EO di Biden sottolineava l'importanza della cooperazione internazionale, incoraggiando l'impegno degli Stati Uniti con alleati e organizzazioni globali per stabilire *standard* di sicurezza e quadri etici comuni per l'IA mentre l'EO di Trump, al contrario, sembra adottare una posizione più unilaterale, affermando la leadership degli Stati Uniti nell'IA senza delineare impegni specifici per la collaborazione internazionale²⁷.

4. *Stargate Project*: la supremazia statunitense nell'intelligenza artificiale attraverso gli investimenti infrastrutturali

Durante il secondo mandato Trump, un secondo significativo filone di scelte strategiche in relazione a cui sembra registrarsi una inversione rispetto alle opzioni privilegiate durante l'amministrazione Biden riguarda le politiche infrastrutturali connesse allo sviluppo dell'AI (la costruzione dei *data center* e i principi che devono essere seguiti nella realizzazione e sviluppo degli stessi).

Contemporaneamente all'adozione dei numerosi EO ed azioni presidenziali citate, il 21 gennaio 2025, il Presidente degli Stati Uniti ha annunciato la creazione di "*Stargate Project*", una *joint venture* tra OpenAI (la società che, tra le tante cose, sviluppa ChatGPT, DALL-E e Sora), SoftBank, Oracle

²⁷ Le riflessioni proposte traggono spunto da P.E. Carrillo - M.J. Mackowski - J.B. Jacobson - A.L. Friel - M. Kirk, *Key Insights on President Trump's New AI Executive Order and Policy & Regulatory Implications*, in *hsquirepattonbogs.com*, 2025.

e la società di investimento MGX, con l'obiettivo di investire fino a 500 miliardi di dollari in infrastrutture per l'intelligenza artificiale negli Stati Uniti entro il 2029²⁸. A tale proposito di investimento ne sono seguiti numerosi altri, tutti meticolosamente rendicontati sulla pagina ufficiale della Casa Bianca²⁹. Il progetto prevede che Masayoshi Son, CEO di SoftBank, assuma il ruolo di presidente dell'impresa.

Scopo del progetto, almeno all'inizio, è la costruzione di "infrastrutture fisiche e virtuali per rafforzare le prossime generazioni di intelligenza artificiale"; dunque, nell'immediato, il progetto si propone la costruzione di enormi *Data center*.

Durante l'annuncio, Trump ha descritto Stargate un'impresa "monumentale" e una "dichiarazione di fiducia verso il potenziale americano". Il progetto prevede un investimento iniziale di 100 miliardi di dollari, destinati a salire a 500 miliardi nei successivi quattro anni. Nelle intenzioni dei loro artefici (oltre a Trump, Sam Altman -CEO di OpenAI-, Larry Ellison -Presidente di Oracle- e Masayoshi Son -CEO di SoftBank-) queste infrastrutture sarebbero finalizzate a garantire la *leadership* americana nell'intelligenza artificiale, creando altresì centinaia di migliaia di posti di lavoro negli Stati Uniti.

Oltre ai principali finanziatori, Microsoft, Nvidia e Arm si sono uniti come partner tecnici, rafforzando la portata strategica del progetto.

I lavori di costruzione dei *Data center* sono immediatamente iniziati in Texas e si prevede un'espansione su scala nazionale ponendo tuttavia numerosi dubbi in tema di consumo energetico necessario.

A questo proposito Trump, contestualmente all'annuncio del progetto Stargate, ha firmato degli ordini esecutivi per dichiarare l'"emergenza energetica nazionale", al fine di garantire una produzione adeguata al fine di sostenere Stargate ed altre iniziative presidenziali³⁰.

Il progetto *Stargate*, chiaramente proposto in termini di concorrenza con il contemporaneo sviluppo in Cina di un modello di AI (*Deepseek*) capace di raggiungere risultati paragonabili a quelli sviluppati dalle tecnologie

²⁸ L'annuncio del Progetto è stato diffuso su innumerevoli testate statunitensi ed internazionali. Fra queste si rinvia al Guardian: *Trump unveils \$500bn Stargate AI project between OpenAI, Oracle and SoftBank*, in *theguardian.com*, 2025; e al sito di Al Jazeera, *Trump Announces \$500bn 'Stargate' venture to build up AI infrastructure*, in *aljazeera.com*, 2025. Il progetto è commentato da G. De Ruvo, *Stargate e l'autoscontro delle stelle rosse*, in *Limes*, 1, 2025, 39 ss.

²⁹ *Trump Effect: A Running List of New U.S. Investment in President Trump's Second Term*, in *whitehouse.gov*.

³⁰ In particolare, nell'EO del 20 gennaio 2025 n. 14156 intitolato "*Declaring a National Energy Emergency*" il Presidente degli Stati Uniti ha proclamato l'emergenza nazionale collegata all'insufficienza della produzione e delle infrastrutture energetiche nel Paese. Sostenendo che le politiche della precedente amministrazione hanno causato una crisi energetica, con alti prezzi dell'energia che danneggiano l'economia e la sicurezza nazionale l'EO denuncia altresì la vulnerabilità del sistema energetico americano alle interferenze di attori stranieri ostili.

L'ordine impone misure di emergenza per aumentare la produzione, il trasporto, la raffinazione e la generazione di energia, con particolare attenzione al Nord-Est e alla costa occidentale degli Stati Uniti incaricando le agenzie federali di utilizzare ogni autorità disponibile per facilitare lo sviluppo di infrastrutture energetiche, accelerare autorizzazioni e regolamenti e garantire un approvvigionamento stabile.

statunitensi a costi decisamente inferiori³¹, non sembra in effetti porsi del tutto in linea con i principi delineati nell'OE di Biden 14141 (*“Advancing United States Leadership in Artificial Intelligence Infrastructure”*) soprattutto con riguardo agli obiettivi di sostenibilità economica (il coinvolgimento nello sviluppo infrastrutturale dell'AI anche di imprese piccole e medie), ecologica e sociale.

Esso, tuttavia, sembra aver a propria volta innescato una reazione europea in tema di investimenti ai fini di costruzione di infrastrutture a servizio dello sviluppo dell'IA in Europa. Nel corso del vertice di Parigi sull'IA la Presidente della Commissione europea Ursula Von der Leyen ha infatti annunciato l'iniziativa *InvestAI* la quale prevede la mobilitazione di 200 miliardi di euro di investimenti per lo sviluppo dell'intelligenza artificiale di matrice europea prevedendo anche un nuovo fondo europeo di 20 miliardi di euro per le *gigafactory* dell'IA³².

5. AI e divergenza regolatoria: l'effetto della seconda amministrazione Trump sull'equilibrio globale

I temi, dunque, che sembrano porsi nell'analisi concernente le divergenze nelle scelte statunitensi al cambiare delle amministrazioni governative federali sono numerosi: i principali attengono alla sostenibilità dello sviluppo dell'AI in tema di tutela dei diritti individuali e allo sviluppo sostenibile delle infrastrutture collegate a tale evoluzione tecnologica.

Le scelte dell'amministrazione democratica sembrano orientate a bilanciare maggiormente i valori in gioco, mentre le scelte repubblicane paiono prevalentemente volte alla deregolamentazione ed alla rimozione delle restrizioni. L'evoluzione delle politiche statunitensi sull'intelligenza artificiale riflette dunque il più ampio dibattito tra regolamentazione e innovazione. Se da un lato l'approccio Trump mira a garantire la supremazia tecnologica americana attraverso la deregolamentazione e investimenti massicci, dall'altro il modello Biden ha tentato di introdurre *standard* di sicurezza ed equità. La seconda amministrazione Trump realizza tuttavia una svolta ancora più netta, con la rimozione di vincoli normativi e l'accelerazione di progetti strategici come *“Stargate Project”*.

Una doverosa riflessione concerne il tema dei riverberi di tali evoluzioni nel contesto europeo.

Rispetto agli Stati Uniti, infatti, l'Unione Europea adotta un approccio più precauzionale, cercando di proporre un *gold standard* globale per l'uso etico

³¹ È stato peraltro osservato come lo sviluppo di un'intelligenza artificiale efficiente ed a basso costo in Cina sia in realtà collegato all'esclusione della Cina dalla possibilità di produrre di chip di qualità cosa che ha costretto il paese a innescare processi di innovazione lavorando sull'ottimizzazione dell'algoritmo e sull'efficienza computazionale. Sul punto cfr. A. Zhang, *Us export control have forced Chinese tech companies to be more innovative*, in *Financial Times* 27 gennaio 2025; S. Rai - N. Purnell, *What is China's DeepSeek and why is it Freaking out the world*, *ibid.* Cfr. anche G. De Ruvo, *Stargate e l'autoscontro delle stelle rosse*, cit.

³² European Commission, *EU launches InvestAI initiative to mobilise €200 billion of investment in artificial intelligence*, 11 febbraio 2025, in *ec.europa.eu*.

dell'intelligenza artificiale, ponendo maggiore enfasi sulla protezione degli utenti e sulla regolamentazione delle tecnologie emergenti.

I numerosi e corposi Regolamenti europei attinenti all'ecosistema digitale -a partire dal Regolamento Generale sulla Protezione dei Dati (GDPR)³³, passando per i già citati *Digital service Act*, *Digital Market Act* e non ultimo l'*AI Act*- sono stati certamente, almeno all'origine, concepiti quali articolati normativi complessi ma armonizzati che intendevano in qualche modo proporsi come un potenziale punto di riferimento internazionale e che si sperava potessero innescare un effetto Bruxelles³⁴.

Tuttavia, nei più recenti mesi sembra che tale evoluzione non sia destinata ad avverarsi giacché le più attuali decisioni della seconda amministrazione Trump sembrano orientarsi verso un approfondimento della divergenza nell'approccio al digitale (e soprattutto in tema di AI) rispetto al contesto europeo, rischiando addirittura una parziale inversione di tendenza nelle strategie europee su questi temi.

Un esempio si può fare con riguardo al tema della responsabilità.

Ad integrazione del regolamento sull'intelligenza artificiale, nel settembre del 2022, era stata difatti proposta dalla Commissione europea la direttiva UE sulla responsabilità per i sistemi di intelligenza artificiale (IA) mirante ad armonizzare le norme sulla responsabilità civile tra gli Stati membri. Tale direttiva era stata pensata al fine di garantire un equo risarcimento per chiunque maturasse danni causati da sistemi IA e dunque introduceva misure specifiche collegate al tema dell'accertamento della responsabilità civile in relazione all'uso di processi algoritmici automatizzati ed intelligenza artificiale.

Le principali disposizioni della direttiva riguardavano: 1) Un tendenziale alleggerimento dell'onere della prova: la direttiva introduceva difatti una "presunzione di causalità" che era finalizzata a facilitare le vittime nel dimostrare il nesso causale tra il danno subito e il comportamento o l'omissione del fornitore, sviluppatore o utente del sistema IA. Questo si rivelava particolarmente significativo in situazioni in cui la complessità dei sistemi IA rende difficile per i danneggiati fornire prove concrete. 2) Un accesso alle prove non opaco: per garantire trasparenza, era previsto che si potesse richiedere l'accesso alle prove detenute dai fornitori o utenti di sistemi IA ad alto rischio. Questo avrebbe permesso di ottenere informazioni cruciali per sostenere le proprie rivendicazioni in caso di danno. 3) Responsabilità oggettiva per sistemi IA ad alto rischio: la direttiva prevedeva che gli operatori di sistemi IA classificati come "ad alto rischio" siano responsabili dei danni causati, indipendentemente dalla colpa. Questo implicava che, in caso di danno, l'operatore era tenuto al risarcimento anche senza prova di negligenza o dolo. 4) Distinzione tra sistemi ad alto e basso rischio: mentre per i sistemi ad alto rischio si applicava la responsabilità oggettiva, per quelli a basso rischio la responsabilità si basava sulla colpa. Questo significava che, in quest'ultimo caso, la vittima doveva dimostrare che il danno derivava da una negligenza o omissione dell'operatore.

³³ Regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio del 27 aprile 2016 avente a oggetto la protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati. Il Regolamento è entrato in vigore nel maggio 2018.

³⁴ Su cui A. Bradford, *Effetto Bruxelles*, Milano, 2021.

Nonostante la intensa volontà di adottare la detta Direttiva, manifestata in più occasioni da personalità ed organi europei, essa è stata infine ritirata (nel gennaio 2025) in stretta concomitanza con l'*AI Action summit* tenutosi a Parigi nel periodo detto; terzo vertice sull'intelligenza artificiale nel corso del quale si è emblematicamente evidenziata una significativa spaccatura fra l'approccio europeo all'intelligenza artificiale e i più recenti orientamenti statunitensi modificatisi a ridosso dell'insediamento della seconda amministrazione Trump³⁵.

Parte della dottrina non smette di ritenere che sul lungo periodo le scelte europee si riveleranno premiali rispetto a quelle più recentemente adottate negli Stati Uniti³⁶. Difatti, la divergenza fra l'enfasi dei provvedimenti di Trump sulla riduzione degli oneri normativi ed il principio di precauzione che dà priorità alle tutele sociali rispetto alla rapida innovazione prescelto nel contesto UE potrebbe creare attriti tra gli ambienti normativi dei due detti ambiti territoriali, in particolare per le aziende multinazionali le quali si muovono in entrambi i sistemi; dunque, sebbene l'*EU AI Act* possa essere considerato, ad un primo rapido sguardo, una normativa limitante che pone degli ostacoli alla libera innovazione, nondimeno la mancanza di tutele etiche esplicite e di misure di mitigazione del rischio nei provvedimenti adottati da Trump potrebbe, dal proprio canto, indebolire la capacità delle aziende statunitensi di competere nei mercati europei, giacché la conformità ai rigorosi standard dell'*AI Act* è un prerequisito legale per l'accesso al mercato dell'UE.

Inoltre, a livello globale, è da rilevare come nelle giurisdizioni di diversi Paesi, fra cui Canada, Giappone, Regno Unito e Australia sono state promosse specifiche politiche concernenti l'AI, molte delle quali sono più in linea con l'attenzione dell'UE sulla responsabilità e sulle considerazioni etiche che con la posizione pro-innovazione degli Stati Uniti tipica dell'amministrazione Trump³⁷. Ad esempio, l'*Artificial Intelligence and Data Act* del Canada enfatizza la trasparenza e lo sviluppo responsabile, mentre le linee guida sull'AI del Giappone promuovono principi di AI affidabili attraverso il coinvolgimento di più parti interessate. Infine, anche il Regno Unito che ha un approccio meno regolamentato rispetto all'UE, pone comunque un forte accento sulla sicurezza attraverso l'AI Safety Institute³⁸.

Le decisioni dell'amministrazione Trump in tema di AI potrebbe porsi in netto contrasto con gli sforzi per stabilire *standard* globali per la *governance* di tale tecnologia. Pertanto, mentre l'UE, il G7 e altre organizzazioni multilaterali stanno lavorando per allinearsi su principi chiave come trasparenza, equità e sicurezza, l'attenzione unilaterale degli Stati Uniti sulla

³⁵ Sul punto cfr. M. Martorana, *Responsabilità per danni da IA: ecco il nuovo percorso normativo Ue*, in *Agenda Digitale*, 5 marzo 2025. La risposta europea alle politiche dell'amministrazione Trump si sono concretate tuttavia anche in maggiori investimenti nel settore dell'AI. Sul punto cfr. S. Montefiori, *Intelligenza artificiale, la sfida di Vance all'Ue: «Troppe regole uccidono l'industria». La risposta europea: «Fondo da 200 miliardi, sarà il Cern dell'AI»*, in *corriere.it*, 11 febbraio 2025.

³⁶ Cfr. P.E. Carrillo - M.J. Mackowski - J.B. Jacobson - A.L. Friel - M. Kirk, *Key Insights on President Trump's New AI Executive Order and Policy & Regulatory Implications*, cit.

³⁷ Si rinvia alle riflessioni esposte *ibid.*

³⁸ Su cui v. il sito dedicato gov.uk/government/organisations/ai-safety-institute.

deregolamentazione potrebbe limitare la sua influenza nel dare forma a queste norme globali. Inoltre, la svolta dell'amministrazione Trump verso la deregolamentazione rischia di creare la percezione che gli Stati Uniti diano priorità ai guadagni di innovazione a breve termine rispetto alle considerazioni etiche a lungo termine, potenzialmente alienando alleati e partner.

Un'ultima considerazione riguarda il rischio che le scelte provvedimentali di Trump amplino il divario tra i regimi normativi federali e statali dell'AI nello stesso contesto statunitense. Difatti, mentre gli EO presidenziali segnalano uno spostamento federale verso la priorità dell'innovazione riducendo i vincoli normativi, a livello subnazionale alcuni Stati³⁹ hanno adottato delle strategie di supplenza normativa promulgando diverse leggi sull'AI alcune più attente al bilanciamento fra diritti ed economia ed altre più vicine alla più recente impostazione federale. Tale attività normativa a livello statale, in assenza di chiare linee guida federali, corre il rischio di portare ad una maggiore frammentazione normativa, mettendo inevitabilmente in difficoltà le aziende del settore.

Nel complesso, dunque, la più recente policy statunitense in tema di AI riflette un cambiamento fondamentale, dando priorità alla deregolamentazione e all'innovazione del libero mercato e riducendo al contempo la supervisione e le tutele etiche. Tuttavia, questo approccio potrebbe non rivelarsi necessariamente premiale e finire per porre in difficoltà le aziende statunitensi che operano in giurisdizioni con normative sull'AI più severe, come l'UE, il Regno Unito, il Canada e il Giappone, nonché alcuni di quegli Stati, nel medesimo contesto statunitense, che hanno già promulgato i propri regimi normativi sull'AI. Nell'attesa di tali possibili sviluppi sarebbe importante che l'Unione Europea non arretrasse nelle sue scelte valoriali e di bilanciamento fra innovazione tecnologica e tutele dei diritti fondamentali.

³⁹ In Colorado nel 2024 è stato approvato il *Colorado AI Act* (con entrata in vigor prevista per il 2026). Si tratta della prima legge statale USA che regola esplicitamente i sistemi di IA ad alto rischio. Esso, per molti aspetti simile all'*AI Act* europeo nel suo impianto, impone ai fornitori obblighi di valutazione del rischio; richiede audit interni e misure di mitigazione del *bias*; prevede obblighi di notifica all'utente quando è coinvolta una decisione algoritmica significativa; introduce diritti di contestazione per i cittadini. In Illinois sin dal 2008 è in vigore il *Biometric Information Privacy Act* (BIPA). Sebbene non si tratti di una legge sull'IA in senso stretto, tale normativa ha avuto un impatto significativo sulle tecnologie che si basano sul riconoscimento facciale e sull'elaborazione dei dati biometrici. Essa impone il consenso esplicito e informato prima della raccolta di dati biometrici; l'obbligo di conservazione e distruzione sicura dei medesimi dati; sanzioni significative in caso di violazione. Diversamente le proposte normative Texane in tema di AI sembrano più vicina all'impostazione della seconda amministrazione Trump. Alcuni disegni di legge (come HB 2060) promuovono la creazione di commissioni per lo studio dell'IA, ma senza imposizioni regolatorie stringenti. Tale impianto si registra anche in altri Stati, fra cui Florida e Georgia. In ultimo particolarmente attento al tema della regolazione dell'AI sembra lo Stato della California ove sono in vigore alcune normative quali il *California Consumer Privacy Act* (CCPA) e il successivo *California Privacy Rights Act* (CPRA) (in vigore dal 2023) le quali prescrivono disposizioni che finiscono per incidere sui temi legati all'AI (profilazione automatizzata, trasparenza, e diritti degli interessati). Nel 2023, è stato presentato un disegno di legge (SB 313) per regolare l'uso di IA nei contesti ad alto rischio, con particolare attenzione alla non discriminazione algoritmica, ma la legge non è stata ancora approvata. Per un maggior approfondimento sul tema si rinvia a BCLP, *State-by-state AI legislation snapshot*, in [bclp.com](https://www.bclp.com).

Abstract

Il contributo analizza l'evoluzione delle politiche federali statunitensi in materia di intelligenza artificiale, mettendo a confronto l'approccio adottato dall'amministrazione democratica Biden e quello affermatosi con il secondo mandato del Presidente Trump. Il lavoro evidenzia come, a fronte di una comune attenzione alla leadership tecnologica e alla sicurezza nazionale, le due amministrazioni abbiano sviluppato visioni profondamente diverse circa il ruolo della regolazione pubblica. Da un lato, l'amministrazione Biden ha promosso un modello di governance "responsabile" dell'innovazione, fondato su strumenti di soft law, sul bilanciamento tra sviluppo tecnologico e tutela dei diritti fondamentali e su un forte coordinamento interagenzia e multilivello. Dall'altro lato, la seconda amministrazione Trump ha inaugurato una marcata inversione di rotta, orientata alla deregolamentazione, alla riduzione dei vincoli etici e sociali e alla riaffermazione della supremazia tecnologica statunitense attraverso investimenti infrastrutturali su larga scala, come il progetto Stargate. Il saggio esamina inoltre le implicazioni di tale divergenza sul piano internazionale, con particolare riferimento al rapporto con l'Unione europea e al rischio di una crescente frammentazione regolatoria globale nel settore dell'intelligenza artificiale.

This article examines the evolution of U.S. federal policies on artificial intelligence by comparing the approach adopted under the Biden administration with the policy shift introduced during President Trump's second term. While both administrations frame AI as a strategic asset for technological leadership and national security, they articulate markedly different conceptions of public intervention. The Biden administration promoted a model of "responsible governance" of innovation, relying primarily on soft law instruments, interagency coordination, and a careful balancing of technological development with the protection of fundamental rights. By contrast, the second Trump administration initiated a decisive turn toward deregulation, reducing ethical and social safeguards and prioritizing American technological dominance through large-scale infrastructure investments, exemplified by the Stargate Project. The article also explores the global ramifications of this regulatory divergence, focusing in particular on transatlantic relations and the growing risk of fragmentation in international AI governance frameworks.

Keywords

intelligenza artificiale - amministrazione federale statunitense - deregolamentazione - governance tecnologica - diritti fondamentali

Tecnologie emergenti e Stati di emergenza: una comparazione tra Belgio e Israele*

Giacomo Giorgini Pignatiello

Sommario

1. Le sfide dell'intelligenza artificiale allo Stato costituzionale contemporaneo. – 2. Metodologia della ricerca e selezione dei casi di studio – 3. Tecnologie emergenti e trasparenza dell'azione amministrativa negli stati di emergenza – 4. *Contact tracing*, sorveglianza digitale e tutela dei diritti fondamentali nella giurisprudenza degli organi di giustizia costituzionale – 4.1. L'ordinamento belga. – 4.1.1. Pandemia da Covid-19 e diritto costituzionale dell'emergenza in Belgio. – 4.2. Il sistema giuridico israeliano. – 4.2.1. Crisi sanitaria e disciplina costituzionale dell'emergenza in Israele. – 4.2.2. *Contact tracing* e Stato di diritto nell'esperienza israeliana – 5. Alcune considerazioni sull'uso delle tecnologie emergenti da parte dei poteri pubblici durante gli stati di emergenza alla luce della comparazione delle esperienze belga e israeliana.

1. Le sfide dell'intelligenza artificiale allo Stato costituzionale contemporaneo

L'erompere dell'emergenza pandemica ha impresso una forte accelerazione al processo di digitalizzazione delle società contemporanee. L'incessante sviluppo delle tecnologie informatiche ha indubbiamente apportato un importante aiuto ai Governi nella gestione di situazioni di crisi di rilevante impatto sulla collettività. Progressivamente algoritmi e intelligenza artificiale sono cominciati a permeare nelle strutture amministrative dell'ordinamento¹, dando così vita al c.d.

*L'articolo è stato sottoposto, in conformità al regolamento della Rivista, a referaggio "a doppio cieco".

Il presente contributo è realizzato nell'ambito del PRIN 2020 "Decision-Making in the Age of Emergencies. New Paradigms in Recognition and Protection of Rights [DeMa]" – CUP C61B22002230001

¹ D. Restrepo-Amariles, *Algorithmic Decision Systems. Automation and Machine Learning in the Public Administration*, in W. Barfield (ed.), *The Cambridge Handbook of the Law of Algorithms*, Cambridge, 2020, 273 ss.; U. Roehl - J. Cromptvoets, *Inside algorithmic bureaucracy: Disentangling automated decision-making and good administration*, in *Public Policy and Administration*, 40(2), 2023, 322 ss.; S. Schäferling, *Governmental Automated Decision-Making and Human Rights*, Cham, 2023; M. Suksi, *The Rule of Law and Automated Decision-Making Exploring Fundamentals of Algorithmic Governance*, Cham, 2023; E. Carloni, *Transparency Within the Artificial Administration Principles, Paths, Perspectives and Problems*, in *The Italian Journal of Public Law*, 1, 2024, 8 ss.; C. Langer,

“Stato digitale”².

Tale irruenta trasformazione, indotta dalla necessità di evitare i contatti umani durante il periodo pandemico, non è stata tuttavia temperata da una necessaria ponderazione sul possibile ruolo delle tecnologie emergenti come fattori in grado di erodere i principi supremi su cui si è affermato nel corso del XX secolo lo Stato di diritto costituzionale³. La resilienza del costituzionalismo contemporaneo agli *stress* determinati dal contestuale operare di rivoluzione digitale e plurime emergenze rispetto a società in continua trasformazione andrà misurata con la capacità degli ordinamenti di riuscire a mantenere salda l'applicazione dei principi fondamentali che governano i sistemi democratico-costituzionali⁴.

Se, infatti, algoritmi e intelligenza artificiale predittiva e generativa soddisfano principi di efficacia, efficienza ed economicità dell'azione pubblica, cruciali in momenti di emergenza nei quali rapidità e chiarezza della risposta dell'ordinamento giuridico risultano vitali per la sopravvivenza della collettività, possono tuttavia sorgere dubbi in relazione a principi di pari rilievo costituzionale, in particolare quelli della trasparenza e della responsabilità (*accountability*) dell'operato del Governo e, più in generale, dei poteri pubblici. Decisioni assunte con modalità opache e non chiaramente imputabili ad un soggetto responsabile, avendo la capacità di limitare, specialmente negli stati di emergenza, diritti fondamentali della persona, dalla salute alla libertà di circolazione, dalla libertà di riunione al diritto all'istruzione, al lavoro, alla libertà di impresa, rischiano di rompere quel patto di fiducia che lega cittadini e apparati pubblici. La concentrazione di grandi quantità di dati (*big data*⁵) in capo a pochi soggetti, la pervasività che contraddistingue le tecnologie informatiche e della comunicazione (c.d. sorveglianza digitale⁶) e l'elevato grado di complessità tecnica rappresenta-

Decision-making power and responsibility in an automated administration, in *Discover Artificial Intelligence*, 4(59), 2024; C. Ramotti, *A Transparent Decision-Making in the Digital Age*, in *The Italian Journal of Public Law*, 1, 2025, 213 ss.; H.C.H. Hofmann - F. Pflücke (eds.), *Governance of Automated Decision-Making and EU Law*, Oxford, 2024; B. Skotnicka-Zasadzień - R. Wolniak, *Artificial Intelligence in Public Administration*, London, 2025; A. Stanojević - E. Keymolen - J. Goossens, *Public Governance and Emerging Technologies*, Cham, 2025.

² Sul concetto di “Stato digitale” in dottrina cfr. M. Simoncini, *Lo Stato digitale. L'agire provvedimentale dell'amministrazione e le sfide dell'innovazione tecnologica*, in *Rivista Trimestrale di Diritto Pubblico*, 2, 2021, 529 ss.; L. Torchia, *Lo Stato digitale*, Bologna, 2023, *passim*; L. Casini, *Il futuro dello Stato (digitale)*, in *Rivista Trimestrale di Diritto Pubblico*, 2, 2024, 433 ss.

³ Per tutti sul tema cfr. P. Ridola, *Il costituzionalismo e lo Stato costituzionale*, in *Nomos*, 2, 2018.

⁴ Sul concetto di resilienza costituzionale cfr. A. Jakab, *Constitutional Resilience*, in R. Grote - F. Lachenmann - R. Wolfrum (eds.), *Max Planck Encyclopedia of Comparative Constitutional Law*, Oxford, 2021.

⁵ M. Falcone, “*Big data*” e pubbliche amministrazioni nuove prospettive per la funzione conoscitiva pubblica, in *Rivista trimestrale di diritto pubblico*, 3, 2017, 601 ss.; F. Costantino, *Rischi e opportunità del ricorso delle amministrazioni alle predizioni dei big data*, in *Diritto pubblico*, 1, 2019, 43 ss.; F. Musella, *Amministrazione 5.0*, in *Rivista di Digital Politics*, 1, 2021, 95 ss.

⁶ S. Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*, New York, 2020. In letteratura v. altresì: G. Resta, *La sorveglianza elettronica di massa e il conflitto regolatorio USA/UE*, in *Il Diritto dell'Informazione e dell'Informatica*, 31(4-5), 2015, 697 ss.; G. Ziccardi, *Sorveglianza elettronica, data mining e trattamento indiscriminato delle informazioni dei cittadini tra esigenze di sicurezza e diritti di libertà*, in *Ragion pratica*, 1, 2018, 29 ss.; G. Formici, *La digital mass surveillance al vaglio della Corte Europea dei Diritti dell'Uomo: da Zakharov*

no alcuni caratteri che rendono l'uso di algoritmi e dell'intelligenza artificiale da parte dei poteri pubblici un fenomeno contraddistinto da precise peculiarità, meritevoli di approfondimento. La sfiducia dei consociati nel perseguimento da parte del Governo e delle amministrazioni dell'interesse pubblico attraverso l'osservanza dei principi dello Stato di diritto costituzionale può catalizzare, in fasi già di per sé delicate nella vita di un ordinamento giuridico, una rottura dell'ordine costituito, con una involuzione democratica del sistema⁷.

Nell'analisi che qui si propone rileva in particolare l'arretramento della componente umana nella gestione e supervisione del procedimento e del provvedimento amministrativo a fronte del ruolo sempre più preminente dell'automazione tecnologica⁸. Proprio a seguito di un periodo emergenziale, che ha posto in dubbio alcuni assunti che dagli anni '90 del secolo scorso avevano plasmato il rapporto tra Stato e individuo, l'attuale ri-espansione dell'intervento pubblico nell'economia e nella società⁹ apre ineludibili questioni relative alla garanzia dei diritti fondamentali dei consociati, specialmente a fronte della verifica di eventi eccezionali in grado di travolgere i principi fondamentali tanto duramente affermatasi con il paradigma costituzionale post-bellico.

Senza alcuna pretesa di esaustività, stante la natura ancora ampiamente inesplorata di fenomeni ancora in via di affermazione, il presente contributo cerca di porre in evidenza alcuni aspetti critici che coinvolgono le garanzie procedurali di rilevanza costituzionale che dovrebbero assistere la formazione delle decisioni pubbliche negli stati di emergenza a fronte delle tecnologie emergenti a cui l'azione pubblica nella sua dimensione multilivello fa sempre più ricorso¹⁰. Dietro un velo di "neutralità tecnolo-

a Big Brother Watch, in *Federalismi.it*, 23, 2020, 44 ss.; M. Panarari, *Il Leviatano digitale e il neototalitarismo* soft, in *Rivista di Digital Politics*, 2, 2021, 397 ss.; M. Westerlund - D.A. Isabelle - S. Leminen, *The Acceptance of Digital Surveillance in an Age of Big Data*, in *Technology Innovation Management Review*, 11(3), 2021, 32 ss.; M. Nino, *La normalizzazione della sorveglianza di massa nella prassi giurisprudenziale delle Corti di Strasburgo e Lussemburgo: verso il cambio di paradigma del rapporto privacy v. security*, in *Freedom, Security & Justice: European Legal Studies*, 3, 2022, 105 ss.; S. Pietropaoli, *Dalla sorveglianza al controllo: la parabola della governamentalità algoritmica*, in *Rivista italiana di informatica e diritto*, 2, 2024, 25 ss.

⁷ Per quanto riguarda il processo di erosione costituzionale in corso anche nelle democrazie liberali, nella sterminata letteratura cfr. P. Howe, *Eroding Norms and Democratic Deconsolidation*, in *Journal of Democracy*, 28, 2016, 15 ss.; J. Chafetz - D. Pozen, *How Constitutional Norms Break Down*, in *UCLA Law Review*, 65, 2018, 1430 ss.; T.G. Daly, *Democratic Decay: Conceptualising an Emerging Research Field*, in *The Hague Journal of the Rule of Law* 11, 2019, 9 ss.; T. Groppi, *Dal costituzionalismo globale ai nuovi autoritarismi. Sfide per il diritto comparato*, in *Rivista AIC*, 4, 2022, 65 ss.

⁸ Cfr. nella letteratura italiana: R. Cavallo Perin - D.U. Galetta (a cura di), *Il Diritto dell'Amministrazione Pubblica digitale*, Torino, 2025, *passim*.

⁹ G. Amato, *Bentornato Stato, ma*, Bologna, 2022, *passim*.

¹⁰ Sul punto in particolare si rinvia all'idea di *digital constitutionalism*, come limite ai grandi poteri digitali e tutela dei diritti fondamentali nel XXI secolo: G. De Gregorio, *Digital Constitutionalism in Europe*, Cambridge, 2022. Altresì: E. Celeste, *Digital constitutionalism: a new systematic theorization*, in *International Review of Law, Computers & Technology*, 33(1), 2019, 76 ss.; G. De Gregorio - R. Radu, *Digital constitutionalism in the new era of Internet governance*, in *International Journal of Law and Information Technology*, 30(1), 2022, 68 ss.; F. de Abreu Duarte - G. De Gregorio - A. Golia, *Perspectives on digital constitutionalism*, in B. Brožek - O. Kanevskaia - P.

gica”¹¹, algoritmi e intelligenza artificiale presentano infatti “potenzialità distorsive” in grado di minare i principi fondamentali del costituzionalismo contemporaneo. Il rischio è dunque che invece di contribuire ad inverare i principi delle democrazie costituzionali contemporanee, algoritmi e intelligenza artificiale possano, se non correttamente governate, favorire violazioni dei diritti fondamentali e minare le basi giuridiche della convivenza civile.

Nei paragrafi successivi l’analisi si concentrerà, con approccio comparato, su come i principi di trasparenza e responsabilità siano stati declinati nella gestione emergenziale in Belgio e in Israele, mettendo in luce i punti di forza e le fragilità dei rispettivi modelli. In particolare, verranno esaminate le decisioni giurisprudenziali e gli assetti normativi che hanno disciplinato l’uso delle tecnologie emergenti nei due ordinamenti, al fine di comprendere se e in che misura tali strumenti abbiano inciso sul delicato equilibrio tra efficienza amministrativa e tutela dei diritti fondamentali.

2. Metodologia della ricerca e selezione dei casi di studio

L’analisi degli episodici casi di studio emersi in giurisprudenza consente di individuare alcuni *trends* ancora in via di consolidamento in relazione alla garanzia del principio di trasparenza e di responsabilità (intesa nell’accezione di *accountability*) degli organi pubblici nell’ambito dei sempre più frequenti procedimenti amministrativi digitali automatizzati, con particolare riguardo agli stati di emergenza. L’attuale frammentarietà delle decisioni giurisprudenziali rese sul punto non permette di ricondurre ad una ricostruzione sistematica compiuta un novero di affermazioni di principio ancora ampiamente provvisorie. Il dato comparato, tuttavia, presenta l’indubbio vantaggio di poter riflettere in senso critico sui risultati cui alcuni organi di giustizia costituzionale sono pervenuti nel mondo nell’ambito della forma di Stato costituzionale contemporaneo.

Uno specifico *cluster* di sentenze è stato individuato in relazione al tema dell’osservanza del principio di trasparenza nei processi di *contact tracing* durante l’emergenza pandemica. Ci si riferisce alle esperienze di Belgio e Israele. La selezione dei casi di studio è avvenuta sulla base dei prototipi¹². Il Belgio appartiene alla tradizione di diritto amministrativo continentale

Palka (eds.), *Research Handbook on Law and Technology*, Cheltenham, 2023, 315 ss.; E. Celeste, *Digital Constitutionalism. The Role of Internet Bills of Rights*, London, 2023; G. De Gregorio - O. Pollicino - P. Valcke (eds.), *The Oxford Handbook of Digital Constitutionalism*, Oxford, 2024; O. Pollicino, *Costituzionalismo digitale*, Bologna, 2025. Per una critica al concetto *cf.* J.R. Gonçalves Pereira - C. Iglesias Keller, *Digital Constitutionalism: Contradictions of a Loose Concept*, in *Rev. Direito e Práx.*, 13(4), 2022, 2648 ss.; A. Golia, *Critique of digital constitutionalism: Deconstruction and reconstruction from a societal perspective*, in *Global Constitutionalism*, 13(3), 2024, 488 ss.; P. Terzis, *Against digital constitutionalism*, in *European Law Open*, 3(2), 2024, 336 ss.

¹¹ G. Carullo, *Principio di neutralità tecnologica e progettazione dei sistemi informatici della pubblica amministrazione*, in *Cyberspazio e diritto*, 21(1), 2020, 33 ss.

¹² R. Hirschl, *The Question of Case Selection in Comparative Constitutional Law*, in *The American Journal of Comparative Law*, 53(1), 2005, 125 ss.

di stampo francese¹³, mentre Israele può ricondursi ad un modello ascrivibile sostanzialmente alla tradizione di *common law*¹⁴.

La selezione dei casi non ha solo un valore descrittivo, ma anche paradigmatico: il Belgio rappresenta un modello multilivello, strettamente ancorato al diritto europeo e al principio di legalità formale; Israele, invece, incarna un approccio securitario, in cui la salute pubblica viene immediatamente assimilata alla sicurezza nazionale. Entrambi i casi offrono quindi prospettive opposte, utili a valutare la capacità di resilienza del costituzionalismo di fronte alla digitalizzazione emergenziale. Si tratta, naturalmente, di una comparazione qualitativa, che non pretende di esaurire il fenomeno, ma mira a identificare linee di tendenza comuni e divergenze strutturali.

Il senso di questa prima ricognizione è quello di delineare in chiave critica alcuni primi principi applicati dalla giurisprudenza costituzionale ai processi di digitalizzazione dei procedimenti amministrativi nello Stato contemporaneo. Stante l'esistenza di una già nutrita letteratura sul punto non verrà trattata l'esperienza italiana¹⁵.

¹³ Per una ricognizione delle metodologie di ricerca nel diritto amministrativo comparato cfr. R. Scarciglia, *Reconsidering Comparative Methodology in Administrative Law*, in *Beijing Law Review*, 10, 2019, 1051 ss.

¹⁴ D. Barak-Erez, *Israeli Administrative Law at the Crossroads: Between the English Model and the American Model*, in *Israel Law Review*, 40(1), 2007, 56 ss.

¹⁵ Per quanto riguarda l'Italia nella sterminata letteratura si vedano: E. Autero - A. Castelli, *Intelligenza artificiale e procedimento amministrativo: paura di cadere o voglia di volare?*, in *Cyberspazio e diritto*, 21(3), 2020, 399 ss.; E. Carloni, *AI, algoritmi e pubblica amministrazione in Italia*, in *Revista de los Estudios de Derecho y Ciencia Política*, 30, 2020, 1 ss.; C. Napoli, *Algoritmi, intelligenza artificiale e formazione della volontà pubblica. La decisione amministrativa e quella giudiziaria*, in *Rivista AIC*, 3, 2020, 318 ss.; P. Forte, *Diritto amministrativo e data science. Appunti di Intelligenza Amministrativa Artificiale (AAI)*, in *PA: Persona e Amministrazione*, 1, 2020, 247 ss.; A.G. Orofino - G. Gallone, *L'intelligenza artificiale al servizio delle funzioni amministrative: profili problematici e spunti*, in *Giurisprudenza italiana*, 7, 2020, 1738 ss.; F. Romano, *Intelligenza Artificiale e amministrazioni pubbliche: tra passato e presente*, in *Cyberspazio e diritto*, 21(1), 2020, 69 ss.; G. Carullo, *Decisione amministrativa e intelligenza artificiale*, in *Il Diritto dell'Informazione e dell'Informatica*, 3, 2021, 431 ss.; R. Cavallo Perin, *L'amministrazione pubblica italiana tra rivoluzione liberale, repubblicana e l'avvento dell'intelligenza artificiale*, in *Aa.Vv., L'amministrazione nell'assetto costituzionale dei poteri pubblici*, Torino, 2021, 9 ss.; A. Lalli (a cura di), *L'amministrazione pubblica nell'era digitale*, Torino, 2022; B. Marchetti, *L'impiego di sistemi di intelligenza artificiale nelle pubbliche amministrazioni italiane: prove generali*, in *BioLaw Journal*, 2, 2022, 489 ss.; N. Rangone, *Intelligenza artificiale e pubbliche amministrazioni: affrontare i numerosi rischi per trarne tutti i vantaggi*, in *BioLaw Journal*, 2, 2022 473 ss.; S.B. Greci, *Le applicazioni di Intelligenza artificiale a supporto dell'automazione del procedimento amministrativo*, in *Rivista italiana di informatica e diritto*, 1, 2024, 217 ss.; E. Carloni, *Dalla legalità algoritmica alla legalità (dell'amministrazione) artificiale. Premesse ad uno studio*, in *Rivista italiana di informatica e diritto*, 2, 2024, 451 ss.; M.C. Cavallaro, *Intelligenza artificiale, interesse pubblico e complessità della scelta amministrativa*, in *Amministrativamente*, 3, 2024, 936 ss.; I. Forgione, *La prospettiva dell'intelligenza artificiale nel procedimento e nell'attività amministrativa*, in *Rivista italiana di informatica e diritto*, 2, 2024, 505 ss.; M. Manna - D. Testa, *Digitalizzazione della pubblica amministrazione e innovazione Territoriale: quo vadis?: la necessità di nuove sintesi per un binomio in (lento) divenire*, in *Munus*, 1, 2024, 141 ss.; R. Cavallo Perin - D.U. Galetta (a cura di), *Il Diritto dell'Amministrazione Pubblica digitale*, Torino, 2025; S. Del Gatto, *Nuove tecnologie, pubblica amministrazione e rapporto con gli amministrati. Brevi note sulle tendenze in atto*, in *Diritto pubblico*, 1, 2025, 155 ss.

3. Tecnologie emergenti e trasparenza dell'azione amministrativa negli stati di emergenza

I sistemi informatici basati sull'uso di algoritmi e di intelligenza artificiale sono notoriamente resistenti alle richieste di controllo esterno. I loro processi non solo sono ammantati di complessità tecnica e opacità¹⁶, ma sono anche spesso protetti da ampie formule di segretezza e diritti di riservatezza¹⁷. Da qui il sorgere in letteratura di molteplici espressioni diverse quali ad esempio: capitalismo di sorveglianza¹⁸, autoritarismo digitale¹⁹, tecno-nazionalismo²⁰, tecno-populismo²¹ e tecno-paternalismo²². Con tale varietà di termini ci si riferisce proprio alla caratteristica presentata in generale dalle nuove tecnologie che combina contestualmente, da una parte, profili di efficienza e tendenziale "oggettività", dall'altra, tendenze distorsive occulte e opacità dei processi decisionali. Accanto agli innegabili vantaggi, si cerca quindi di evidenziare quegli aspetti delle tecnologie, che rischiano di erodere la fiducia dei consociati nel concreto operare delle istituzioni democratiche, a fronte di meccanismi decisori incomprensibili per i cittadini destinatari degli interventi pubblici. L'opacità, infatti, rendendo imperscrutabile l'operato dei pubblici poteri rompe quel rapporto di fiducia che in una democrazia costituzionale dovrebbe legittimare l'azione pubblica²³. Tale rottura può generare una degradazione dell'ordine

¹⁶ F. Pasquale, *The Black Box Society: The Secret Algorithms That Control Money and Information*, Cambridge (Mass.), 2015.

¹⁷ M.J. Ryan, *Secret Algorithms, IP Rights, and the Public Interest*, in *Nevada Law Journal*, 21(1), 2020-2021, 61 ss.; K. Foss-Solbrekk, *Searchlights across the black box: Trade secrecy versus access to information*, in *Computer Law & Security Review*, 50, 2023; H. Sun, *The Right to Know Social Media Algorithms*, in *Harvard Law & Policy Review*, 18(1), 2023, 1 ss.

¹⁸ S. Zuboff, *The Age of Surveillance Capitalism*, cit.

¹⁹ T. Roberts - M. Oosterom, *Digital authoritarianism: a systematic literature review*, in *Information Technology for Development*, 2024, 1 ss.

²⁰ S. Nakayama, *Techno-Nationalism versus Techno-Globalism*, in *East Asian Science, Technology and Society: An International Journal*, 6(1), 2012, 9 ss.

²¹ P. Bloom - A. Sancino, *Disruptive democracy: the clash between techno-populism and techno-democracy*, California, 2019.

²² W.-C. Chang, *Asian Values, Confucianism, and Illiberal Constitutions*, in A. Sajó - R. Uitz - S. Holmes (eds.), *Routledge Handbook of Illiberalism*, New York, 2021, 82 ss.

²³ Sul principio costituzionale della trasparenza amministrativa nella letteratura italiana si vedano solo esemplificativamente: F. Merloni (a cura di), *La trasparenza amministrativa*, Milano, 2008; E. Carloni, *La casa di vetro e le riforme: modelli e paradossi della trasparenza amministrativa*, in *Diritto pubblico*, 3, 2009, 779 ss.; A.G. Orofino, *Profili giuridici della trasparenza amministrativa*, Bari, 2013; L. Califano - C. Colapietro (a cura di), *Le nuove frontiere della trasparenza nella dimensione costituzionale*, Napoli, 2014; F. Cardarelli, *Amministrazione digitale, trasparenza e principio di legalità*, in *Il diritto dell'informazione e dell'informatica*, 2, 2015, 227 ss.; C. Colapietro, *La terza generazione della trasparenza amministrativa: dall'accesso documentale all'accesso generalizzato, passando per l'accesso civico*, Napoli, 2016; F. de Leonardis, *Tra obblighi di trasparenza e diritto alla riservatezza: verso una trasparenza di tipo funzionale*, in *Giornale di storia costituzionale*, 31(1), 2016, 175 ss.; G.M. Salerno, *Verso un'amministrazione trasparente: principi costituzionali e regole legislative*, in *Giornale di storia costituzionale*, 31(1), 2016, 97 ss.; S. Foà, *La nuova trasparenza amministrativa*, in *Diritto Amministrativo*, 1, 2017, 65 ss.; A. Lazzaro, *Trasparenza e prevenzione della cattiva amministrazione*, Milano, 2017; C.P. Santacroce, *Conquiste e arretramenti nella trasparenza amministrativa che cambia*, in *Diritto e società*, 3, 2017, 577 ss.; A. Celotto, *La trasparenza e l'accesso ai documenti amministrativi*,

costituzionale.

La digitalizzazione dei processi decisionali delle amministrazioni pubbliche ed in particolare il ricorso sempre maggiore ad algoritmi e intelligenza artificiale, quale fenomeno che coinvolge trasversalmente il settore pubblico non meno di quello privato, solleva, dunque, preoccupazioni in ordine alla centralizzazione dell'esercizio pervasivo di un potere immune da controlli ed equilibri tradizionali²⁴.

4. Contact tracing, sorveglianza digitale e tutela dei diritti fondamentali nella giurisprudenza degli organi di giustizia costituzionale

4.1. L'ordinamento belga

4.1.1. Pandemia da Covid-19 e diritto costituzionale dell'emergenza in Belgio

Il Belgio è una monarchia parlamentare federale caratterizzata da una forte autonomia delle Regioni e delle Comunità e un elevato decentramento di Province e Municipi, tutti costituzionalmente previsti²⁵. La previsione di un *Comité de concertation/Overlegcomité* (CCO) rappresenta uno snodo istituzionale fondamentale nel coordinamento dei molteplici piani di governo che caratterizzano lo Stato belga²⁶. La Costituzione belga non solo non prevede una specifica clausola dedicata agli stati di emergenza, ma anzi all'art. 187 fa espresso divieto di una sospensione del Testo fondamentale, nella sua interezza o in alcune sue parti²⁷. L'ordinamento belga, tuttavia,

in C. Contessa - R. Greco (a cura di), *L'attività amministrativa e le sue regole*, Piacenza, 2020, 541 ss.; E. Carloni, *Il paradigma trasparenza. Amministrazione, informazione, democrazia*, Bologna, 2022; B. Ponti, *Dieci anni di diritto alla trasparenza amministrativa: caratteri originari, trasformazione e integrazione del modello*, in *Le Istituzioni del Federalismo*, 3, 2022, 635 ss.; T. Altì - C. Barbieri, *La trasparenza amministrativa come strumento di potere e di democrazia*, in *Rivista trimestrale di diritto pubblico*, 2, 2023, 809 ss.; E. Carloni, *La trasparenza amministrativa e gli algoritmi*, in E. Belisario - G. Cassano (a cura di), *Intelligenza artificiale per la pubblica amministrazione*, Pisa, 2023, 219 ss.

²⁴ I. Gontarz, *Judicial Review of Automated Administrative Decision-making: The Role of Administrative Courts in the Evaluation of Unlawful Regimes*, in *ELTE Law Journal*, 1, 2023, 155.

²⁵ Per una panoramica sull'ordinamento costituzionale belga: A. Mastromarino, *Belgio*, Bologna, 2013.

²⁶ Sulla gestione multilivello della risposta pandemica in Belgio v. P. Bursens - P. Popelier - K. Verhoest, *Belgium: Cooperation Out of Necessity and Confirmation of Executive Dominance*, in K. Lynggaard - M.D. Jensen - M. Kluth (eds.), *Governments' Responses to the Covid-19 Pandemic in Europe*, Cham, 2022, 387 ss.; D. Sinardet - J. Pieters, *Nationalism and COVID in Belgium: A Surprisingly United Response in a Divided Federal Country*, in A. Lecours - S. Kerr (eds.), *Multinationalism and Covid-19*, London, 2023, 55 ss.; P. Bursens - P. Popelier - P. Meier, *Belgium's Response to COVID-19. How to manage a pandemic in a competitive federal system?*, in R. Chattopadhyay et al. (eds.), *Federalism and the Response to COVID-19. A Comparative Analysis*, Abingdon, 2022, 39 ss.

²⁷ Nella letteratura italiana cfr. G. Milani, *L'emergenza sanitaria nel diritto pubblico comparato: la risposta del Belgio al Covid-19*, in *Dpce online*, 2, 2020, 1671 ss.; C. Marchese, *La governance multilivello delle emergenze nell'ordinamento belga. Considerazioni a partire dall'emergenza pandemica*, in

presenta una spiccata apertura verso il diritto internazionale e sovranazionale, tanto da considerarsi aderente alla teoria “monista”. Ciò ha avuto un importante ruolo nel periodo pandemico, consentendo limitazioni dei diritti fondamentali dei consociati proprio sulla base del principio di precauzione disegnato all’art. 191 del TFUE. Le misure più importanti per il contrasto dell’emergenza sanitaria sono state assunte dal Governo federale, sulla base della normativa emergenziale pre-pandemica²⁸.

I principali interventi normativi volti a limitare il contagio pandemico sono avvenuti tramite decreti del Ministro dell’Interno, che ha agito sulla base dei poteri derivanti dal combinato disposto della legge sulla protezione civile del 1963, della legge sulle forze di polizia del 1992 e della legge sulla sicurezza civile del 2007. Le leggi del 1963 e del 1992 specificano e definiscono gli obiettivi delle misure che il Ministro dell’Interno può adottare (tutela dell’ordine e della salute pubblica, nonché protezione della popolazione), ma non la natura normativa di tali provvedimenti. In forza del principio di sussidiarietà verticale, il Ministro dell’Interno può agire solo quando le misure adottate a livello locale non sono sufficienti a tutelare adeguatamente la salute pubblica. La legge del 2007, inoltre, attribuisce al Ministro dell’Interno il potere di limitare la libertà di movimento dei cittadini quando circostanze pericolose lo richiedono²⁹.

Dopo una prima fase particolarmente critica (febbraio-marzo 2020), i successivi decreti ministeriali che hanno istituito o attenuato il lockdown sono stati tutti ampiamente discussi all’interno del Consiglio di Sicurezza Nazionale (CSN - fino a settembre 2020) e successivamente dal CCO (da ottobre 2020) prima della loro adozione. Il CSN è responsabile della definizione della politica generale di *intelligence* e sicurezza nel Paese. Inoltre, coordina e determina le priorità dei servizi di *intelligence* e sicurezza. Presieduto dal primo ministro, è composto dai ministri della Giustizia, della Difesa nazionale, degli Interni e degli Esteri, nonché da tutti gli (altri) vice primi ministri. Durante la crisi sanitaria, comprendeva anche i ministri-presidenti delle entità federate. Il passaggio alla discussione in seno al CCO ha testimoniato il desiderio delle Regioni e delle Comunità di avere una maggiore influenza nello sviluppo della strategia belga contro la pandemia: nell’OCC, infatti, il Governo federale e gli esecutivi regionali e comunitari sono posti su un piano di parità.

A livello federale, il 27 marzo 2020 sono state adottate due leggi sui poteri speciali per fronteggiare speditamente ed efficacemente l’emergenza Covid-19, con un ampio sostegno parlamentare. Il Parlamento ha concesso poteri speciali al Governo federale per tre mesi (30 marzo – 30 giugno

Dpce online, 3, 2023, 2693 ss.

²⁸ Regio decreto che stabilisce il piano di emergenza per eventi e situazioni di crisi che richiedono un coordinamento o una gestione a livello nazionale (31 gennaio 2003) e Regio decreto sulla pianificazione e sulla gestione delle emergenze a livello comunale e provinciale (22 maggio 2019).

²⁹ M. Verrdussen, *La Constitution belge face à la pandémie de Covid-19*, in *Confluence des droits*, 9, 2020; P. Bouvier, *L’État de droit et la crise sanitaire en Belgique*, in D. Le Prado (ed.), *L’État de droit et la crise sanitaire*, Paris, 2022, 103 ss.; P. Popelier - C. Van de Heyning - S. Van Drooghenbroeck, *National Report on Belgium*, in A. Vedaschi (ed.), *Governmental Policies to Fight Pandemic*, Leiden, 2024, 137 ss.

2020). Le leggi sui poteri speciali hanno previsto che il Parlamento dovesse confermare i decreti adottati dal Governo entro un anno dalla loro entrata in vigore. In caso contrario, i decreti avrebbero perso la loro validità. Le leggi sui poteri speciali stabilivano inoltre che il Governo avrebbe potuto utilizzare i propri poteri speciali solo per combattere la pandemia da Covid-19 e alleviarne le conseguenze, specificando in modo dettagliato gli obiettivi che potevano essere perseguiti dal Governo. Tra questi, in particolare, si rinvenivano: garantire la sicurezza degli approvvigionamenti, fornire sostegno alle persone e alle associazioni che hanno sofferto a causa della pandemia, assicurare la continuità e la stabilità dell'economia, adattare le scadenze stabilite dalla legge o assicurare la continuità del sistema giudiziario³⁰.

Tutte le norme adottate durante lo Stato di emergenza potevano essere sottoposte al vaglio della giurisdizione al fine di verificare l'osservanza principio di legalità e il rispetto dei diritti fondamentali dei consociati³¹. Il contenzioso, sebbene non abbia condotto a particolari censure da parte della giurisdizione, è stato tuttavia piuttosto intenso, sintomo comunque del buon funzionamento del potere giudiziario anche in una situazione emergenziale. I regolamenti regi, governativi e ministeriali, inoltre, sono normalmente sottoposti al controllo preventivo di legittimità da parte della Sezione per la Legislazione del Consiglio di Stato³².

Per quanto riguarda il Parlamento, è importante ricordare che in qualunque momento quest'ultimo può paralizzare pro-futuro l'efficacia dei decreti governativi con i quali è stata regolata la pandemia, abrogando la fonte legislativa con la quale veniva attribuito al Governo il relativo potere normativo³³. Nel contesto della crisi da Covid-19, diversi parlamenti in Belgio, tra cui quello federale, hanno creato comitati parlamentari *ad hoc* per confermare i decreti governativi adottati sulla base delle leggi sui poteri speciali e per condurre inchieste sulla gestione della crisi da parte dei Governi locali e federale. Tali comitati parlamentari hanno contribuito in maniera significativa ad aumentare il controllo parlamentare sull'uso dei poteri speciali da parte degli esecutivi. I parlamenti in Belgio hanno continuato ad operare

³⁰ P. Popelier, *COVID-19 legislation in Belgium at the crossroads of a political and a health crisis*, in *The Theory and Practice of Legislation*, 8(1–2), 2020, 131 ss.; E. Slautsky et al., *Belgium: Legal Response to Covid-19*, in J. King - O. Ferraz (eds.), *The Oxford Compendium of National Legal Responses to Covid-19*, Oxford, 2022.

³¹ Sulla giurisprudenza costituzionale belga nel periodo pandemico *cf.* L. Sciannella, *Stati di emergenza e tenuta degli ordinamenti democratici. La giurisprudenza della "Cour constitutionnelle" belga sulla legislazione pandemica*, in *DPCE*, 2, 2025, 319 ss.

³² Sul ruolo della Sezione per la Legislazione del Consiglio di Stato in Belgio *cf.* E. Thibaut, *Le Conseil d'État de Belgique dans ses fonctions consultatives et législatives*, in *Parliaments, Estates and Representation*, 43(1), 2023, 19 ss.

³³ Per una disamina dell'attività parlamentare nel periodo pandemico in Belgio *v.* F. Bouhon et al., *L'État belge face à la pandémie de Covid-19: esquisse d'un régime d'exception*, in *Courrier hebdomadaire du CRISP*, 1, 2020, 5 ss.; C. Fallon - A. Thiry - S. Brunet, *Planification d'urgence et gestion de crise sanitaire. La Belgique face à la pandémie de Covid-19*, in *Courrier hebdomadaire du CRISP*, 8, 2020, 5 ss.; A. Jousten - C. Behrendt, *Fonctionnement des parlements belges en période de confinement et de distanciation sociale*, in F. Bouhon - E. Slautsky - S. Wattier (eds.), *Le droit public belge face à la crise du COVID-19*, Bruxelles, 2022, 225 ss.; V. Lefebvre, *La démocratie face aux «crises» globales: la Belgique entre unité et division*, in *Courrier hebdomadaire du CRISP*, 34–35, 2024, 5 ss.

durante tutta la pandemia, occupandosi quasi interamente della discussione delle norme relative all'emergenza pandemica. L'attività parlamentare è stata garantita attraverso il ricorso al distanziamento sociale, o attraverso modalità ibride di partecipazione ai lavori, quando non interamente svolti tramite videoconferenza.

4.1.2. La Corte costituzionale belga e la tutela del principio di trasparenza nel sistema di tracciamento dei contagi

In Belgio, diversi livelli di governo sono stati coinvolti nel rintracciamento manuale e digitale delle persone infette da Covid-19, delle persone sospette di infezione e dei loro contatti. Il governo federale e diverse entità federate regionali hanno stipulato un accordo di cooperazione per organizzare la ricerca dei contatti. L'accordo di cooperazione prevedeva la creazione di diversi *database*, tra cui "Sciensano", gestiti da enti e livelli di governo differenti³⁴.

Alcune associazioni e deputati di lingua tedesca hanno impugnato l'accordo di cooperazione davanti alla Corte costituzionale, che ha emesso un'importante sentenza (n. 110 del 2022)³⁵.

La decisione n. 110/2022 della Corte costituzionale belga è intervenuta in modo diretto al cuore delle questioni costituzionali poste dalla gestione della pandemia da Covid-19 attraverso l'uso da parte degli esecutivi di strumenti normativi e amministrativi fondati sulla raccolta, il trattamento e la circolazione di dati personali e sanitari su larga scala, collocandosi pienamente all'interno di quella zona grigia in cui l'emergenza sanitaria, pur non formalizzata come stato di emergenza costituzionale in senso stretto, produce effetti strutturalmente analoghi a quelli di un regime eccezionale, incidendo in modo profondo e duraturo sull'esercizio dei diritti fondamentali e sulla configurazione dei poteri pubblici. La Corte è stata chiamata a pronunciarsi sulla legittimità costituzionale di un complesso normativo che istituisce e disciplina una pluralità di banche dati Covid-19, destinate al tracciamento dei contagi, alla sorveglianza epidemiologica e al coordinamento informativo tra autorità federali ed entità federate, e sin dalle prime battute chiarisce che l'oggetto del giudizio non è la scelta politica di contrastare la pandemia mediante strumenti digitali, ma la compatibilità delle modalità giuridiche prescelte con i parametri costituzionali della tutela della vita privata e della protezione dei dati personali, letti in connessione con i principi di legalità, proporzionalità e responsabilità.

La Corte prende le mosse dalla constatazione che il legislatore, nel contesto di una crisi sanitaria senza precedenti, ha fatto ricorso a strumenti normativi adottati in tempi rapidi, che hanno consentito la creazione di in-

³⁴ E. Degrave, *Les citoyens contrôlés via leurs données Covid? Le «datamatching» et le «datamining» utilisés par l'État*, in *Journal des Tribunaux*, 2021, 125 ss.; E. Degrave, *Traçage, fichage, profilage: la vie privée touchée par le Covid*, in F. Bouhon - E. Slautsky - S. Wattier (eds.), *Le droit public belge face à la crise du COVID-19: quelles leçons pour l'avenir?*, Bruxelles, 2022, 785 ss.

³⁵ Corte costituzionale belga, decisione n. 110/2022/ECLI:BE:GHCC:2022:ARR.110.

infrastrutture informative complesse e interconnesse, capaci di raccogliere e trattare dati particolarmente sensibili relativi allo stato di salute, ai contatti personali e ai comportamenti degli individui, e riconosce che tale scelta risponde a una finalità di interesse generale di eccezionale rilevanza, ossia la tutela della salute pubblica e il contenimento della diffusione del virus, idonea in linea di principio a giustificare interferenze anche significative con il diritto al rispetto della vita privata. Al tempo stesso, tuttavia, la Corte ha implicitamente rifiutato una lettura “emergenzialista” che consideri tali interferenze come automaticamente legittime o sottratte al controllo di costituzionalità, affermando che la gravità della situazione sanitaria non sospende né attenua le esigenze di legalità e di garanzia, che presidiano la limitazione dei diritti fondamentali, ma semmai le rende più stringenti, soprattutto quando le misure adottate si fondano su strumenti tecnologici che amplificano la capacità del potere pubblico di incidere sulla sfera individuale. In questo senso, la Corte ha chiarito che l’istituzione di banche dati e sistemi di tracciamento non è in sé incompatibile con i principi costituzionali, purché tali strumenti siano previsti da una base legale sufficientemente chiara e accessibile, perseguano una finalità legittima e rispettino il principio di proporzionalità, inteso non solo come adeguatezza e necessità delle misure rispetto allo scopo perseguito, ma anche come esigenza di delimitazione precisa dell’ambito e delle modalità dell’interferenza.

È proprio su questo terreno che il giudizio della Corte si fa particolarmente analitico, poiché essa individua nella qualità della disciplina normativa – più che nella scelta di utilizzare strumenti digitali in quanto tali – il punto critico del bilanciamento tra tutela della salute e diritti fondamentali, osservando che, in presenza di trattamenti di dati personali su larga scala, la legge deve definire con particolare rigore le condizioni di liceità, le responsabilità istituzionali e i limiti temporali delle misure, al fine di evitare che l’eccezione pandemica si traduca in una normalizzazione di forme invasive di raccolta e conservazione dei dati. Sotto questo profilo, la Corte concentra una parte rilevante del proprio scrutinio sulla questione della responsabilità e della supervisione delle banche dati Covid-19, rilevando che la normativa impugnata non individua in modo sufficientemente chiaro un’autorità federale incaricata del coordinamento e del controllo complessivo dei trattamenti, ma distribuisce competenze e funzioni tra diversi livelli di governo senza assicurare un centro di imputazione univoco della responsabilità, con il risultato di indebolire il principio di accountability che costituisce un pilastro tanto della tutela costituzionale della vita privata quanto del diritto europeo della protezione dei dati. La Corte osserva che tale frammentazione non rappresenta un problema meramente organizzativo, ma incide direttamente sulla sostanza dei diritti, poiché rende più difficile per gli interessati comprendere chi sia responsabile del trattamento dei loro dati, a chi rivolgersi per esercitare i propri diritti e chi risponda in caso di uso improprio o eccessivo delle informazioni raccolte, e sottolinea che, in un contesto emergenziale caratterizzato da urgenza e rapidità decisionale, l’assenza di una chiara catena di responsabilità rischia di tradursi in una riduzione effettiva delle garanzie.

Accanto a questo profilo, la Corte individua una seconda criticità di rilievo costituzionale nella disciplina della conservazione dei dati personali,

rilevando che la normativa contestata non prevede limiti temporali sufficientemente precisi e stringenti alla durata del trattamento, ma si limita a richiamare genericamente la finalità di gestione della crisi sanitaria, lasciando così aperta la possibilità che i dati siano conservati per un periodo indeterminato o comunque eccedente rispetto a quanto strettamente necessario. La Corte insiste sul fatto che il principio di limitazione temporale del trattamento non è un elemento accessorio della tutela della privacy, ma una garanzia essenziale contro il rischio che misure adottate in un contesto eccezionale si protraggano oltre la durata dell'emergenza, trasformandosi in infrastrutture permanenti di raccolta e conservazione dei dati, e osserva che, in assenza di termini chiari fissati dalla legge, la temporaneità delle misure resta affidata alla discrezionalità dell'amministrazione, con una compressione indebita del principio di legalità. Nelle proprie argomentazioni, la Corte collega esplicitamente questo profilo alla natura stessa delle tecnologie utilizzate, evidenziando come i sistemi digitali di gestione dei dati rendano particolarmente agevole la conservazione prolungata delle informazioni e, di conseguenza, richiedano un intervento normativo ancora più puntuale per evitare derive funzionali o utilizzi successivi non originariamente previsti.

La Corte affronta inoltre il tema della condivisione dei dati e delle modalità di accesso da parte di soggetti terzi, rilevando che alcune disposizioni attribuiscono a organismi tecnici poteri ampi di autorizzazione alla comunicazione dei dati senza un sufficiente ancoraggio a criteri normativi dettagliati e senza un adeguato controllo parlamentare, e osserva che tale configurazione istituzionale solleva problemi di trasparenza e di controllo democratico, soprattutto in un contesto in cui la gestione dell'emergenza sanitaria ha comportato una significativa espansione del potere amministrativo.

Nel valutare complessivamente il bilanciamento operato dal legislatore, la Corte ribadisce che la protezione della salute pubblica costituisce un obiettivo di primario rilievo costituzionale, ma afferma con altrettanta chiarezza che essa non può essere perseguita a costo di un' indefinita compressione del diritto alla vita privata e alla protezione dei dati personali, e che il ricorso a strumenti tecnologici avanzati non attenua, ma anzi rafforza, l'esigenza di garanzie giuridiche precise, controllabili e temporalmente delimitate. Il controllo di costituzionalità esercitato dalla Corte si configura così come un controllo sostanziale sulla qualità della legge, volto a verificare se il legislatore abbia effettivamente assunto su di sé la responsabilità di delimitare l'eccezione, o se abbia invece trasferito tale responsabilità all'amministrazione, sfruttando la flessibilità delle tecnologie digitali per estendere nel tempo e nello spazio l'interferenza sui diritti. Sul piano degli effetti, la Corte opta per una soluzione calibrata, annullando le disposizioni ritenute incompatibili con la Costituzione ma prevedendo un regime transitorio che consenta al legislatore di intervenire per colmare le lacune individuate, al fine di evitare un vuoto normativo che potrebbe compromettere la gestione della crisi sanitaria, e questa scelta conferma che l'obiettivo della decisione non è quello di delegittimare l'uso di strumenti digitali nella governance dell'emergenza, bensì di ricondurli entro un quadro di legalità costituzionale capace di prevenire la normalizzazione dell'eccezione.

Complessivamente, la decisione n. 110/2022 offre quindi una ricostruzione particolarmente significativa del rapporto tra stato di emergenza, tecnologie digitali e tutela dei diritti fondamentali, mostrando come l'ordinamento costituzionale belga affronti la crisi pandemica non attraverso una sospensione implicita delle garanzie, ma mediante un controllo rigoroso delle modalità con cui l'emergenza viene giuridicamente strutturata, e affermando che la resilienza costituzionale non risiede nel rifiuto della tecnologia, ma nella capacità di assoggettarla a regole chiare, responsabilità definite e limiti temporali stringenti, tali da impedire che strumenti nati per fronteggiare una situazione straordinaria si trasformino in dispositivi ordinari di compressione dei diritti fondamentali.

4.2. Il sistema giuridico israeliano

4.2.1. Crisi sanitaria e disciplina costituzionale dell'emergenza in Israele

Sin dalla propria fondazione nel 1948, nell'ordinamento israeliano è in vigore uno stato di emergenza permanente, dovuto ai rapporti particolarmente tesi con gli Stati confinanti³⁶. Lo stato di emergenza è attualmente disciplinato nel Testo fondamentale israeliano (*Basic law*). La Knesset, il Parlamento monocamerale israeliano, autorizza regolarmente con cadenza annuale la proroga dello stato di emergenza³⁷.

La Legge fondamentale israeliana consente al Governo durante lo stato di emergenza di adottare regolamenti di emergenza: "Per la difesa dello Stato, per garantire la sicurezza pubblica e per assicurare l'approvvigionamento di beni e servizi essenziali". Il regolamento governativo deve essere presentato alla Commissione Affari Esteri e Sicurezza presso la Knesset e perde la propria validità trascorsi tre mesi dalla propria emanazione, a meno che la Knesset non lo abroghi prima. La Knesset, tuttavia, può anche prolungarne la durata, adottando in tal senso un'apposita legge.

Nei primi mesi della pandemia da Coronavirus, il Governo è ricorso all'adozione di regolamenti emergenziali, come mai aveva fatto in precedenza, neanche durante le fasi belliche più intense³⁸. I regolamenti emergenziali hanno investito, come in molti altri Paesi del mondo, ambiti molto delicati

³⁶ A. Mizock, *The Legality of the Fifty-Two Year State of Emergency in Israel Contents*, in U.C. Davis *Journal of International Law & Policy*, 7(2), 2001, 223 ss.; B. Bracha, *Checks and Balances in a Protracted State of Emergency - The Case of Israel*, in *Israel Yearbook on Human Rights*, Leiden, 2003, 123 ss.; S. Navot, *Emergency as a State of Mind—The Case of Israel*, in P. Auriel - O. Beaud - C. Wellman (eds.), *The Rule of Crisis. Terrorism, Emergency Legislation and the Rule of Law*, Cham, 2018, 185 ss.

³⁷ Per una ricostruzione sull'attuale stato di crisi politico-costituzionale di Israele v. D. Kretzmer, *Israel's Constitutional Crisis*, in *Quaderni costituzionali*, 1, 2023, 178 ss.

³⁸ E. Albin et al., *Israel: Legal Response to COVID-19*, in J. King - O. Ferraz (eds.), *The Oxford Compendium of National Legal Responses to Covid-19*, Oxford, 2022; D. Shtaubert - G. Harari-Heit - I. Bar-Siman-Tov, *Sunset Legislation in Israel during the COVID-19 Pandemic*, in *Taiwan Journal of Democracy*, 19(2), 2023, 77 ss.; I. Bar-Siman-Tov et al., *Sunsets and Emergencies: Temporary Covid Laws in Israel*, 3 April 2024, in *ssrn.com*.

della vita delle persone: restrizioni alla libertà di circolazione, distanziamento sociale, tracciamento dei contagi. Un così massiccio ricorso ai regolamenti emergenziali da parte del Governo ha dato luogo a diverse critiche in dottrina³⁹. È stato osservato, infatti, che si sarebbe comunque potuto seguire un iter legislativo ordinario nella Knesset, avendo il Parlamento continuato a lavorare anche durante l'emergenza pandemica. Da alcuni è stato altresì ritenuto, adottando un'interpretazione restrittiva, che i regolamenti emergenziali sarebbero utilizzabili dal Governo solo quando è a rischio la sicurezza nazionale del Paese. Ancora, vi è chi ha sostenuto che in una prima fase il ricorso ai regolamenti emergenziali sarebbe stato effettivamente legittimo, in quanto, a seguito delle elezioni politiche nazionali svoltesi il 2 marzo 2020, la Knesset non era entrata ancora nel pieno delle sue funzioni. Una volta costituite e divenute operative le commissioni parlamentari presso la Knesset, tuttavia, il Governo si sarebbe dovuto esimere dall'adozione di nuovi regolamenti emergenziali, ben potendo operare attraverso l'iter legislativo ordinario. Pur non trovando un'apposita censura in sede giurisdizionale⁴⁰, tali posizioni hanno convinto in breve tempo la Knesset a riappropriarsi della propria funzione normativa. Nel luglio 2020 è stata così adottata la Legge sul Coronavirus.

Il secondo strumento legislativo pre-Covid-19 più diffuso, utilizzato per misure come la quarantena, è stato quello delle ordinanze del Ministero della Sanità, che trovano la propria fonte sulla produzione nella *Public Health Ordinance 1940* (PHO), un atto legislativo risalente all'epoca del "Mandato britannico", che conferisce al Ministero della Sanità il poter di adottare vari provvedimenti sanitari per contenere le malattie infettive. La prima serie di disposizioni, contenute nel PHO 1940, enumerate sotto la rubrica "poteri di emergenza", si applica solo quando viene dichiarata l'esistenza di un grave pericolo di malattia contagiosa. In questa situazione, il Ministero della Salute dispone di poteri speciali per prevenire la diffusione della malattia, oltre a quelli generali previsti dalla legge. La seconda serie di disposizioni si applica, invece, solo alle malattie indicate in apposito elenco allegato al PHO 1940, in cui sono specificate le malattie contagiose di rilevanza internazionale, che richiedono una notifica immediata ai sensi del Regolamento sanitario internazionale (2005) della WHO. Il 27 gennaio 2020 il Ministro della Sanità ha provveduto ad inserire anche il Covid-19 tra tali malattie infettive.

Il 23 luglio 2020, la Knesset ha approvato la legge sulle autorità speciali per affrontare il nuovo coronavirus (disposizione temporanea) 2020 (c.d. "Legge sul Corona")⁴¹, che consente al Governo (invece che alla Knesset) di dichiarare: "Una situazione di emergenza a causa del coronavirus" se è convinto che vi sia un rischio reale di diffusione del virus e di danni signi-

³⁹ M. Maor - R. Sulitzeanu-Kenan - D. Chinitz, *When COVID-19, constitutional crisis, and political deadlock meet: the Israeli case from a disproportionate policy perspective*, in *Policy and Society*, 39(3), 2020, 442 ss.

⁴⁰ *Adalah – The Legal Center for Arab Minority Rights in Israel v. Prime Minister Netanyahu* [2020] 2399/20 HCJ (HCJ), para 5.

⁴¹ Per una disamina sulla legislazione emergenziale adottata da Israele durante la pandemia da covid-19 v. I. Bar-Siman-Tov et al., *Temporary COVID laws*, in F. Fagan - S. Levmo (eds.), *Research Handbook on Law and Time*, Cheltenham, 2025, 275 ss.

ficativi alla salute pubblica. La Legge sul Corona è stata preceduta da una cosiddetta “mini-legge sul Corona”, che è stata approvata nel giugno 2020, mentre era in corso l’iter legislativo per l’adozione della “Grande-Legge sul Corona” (come è stata soprannominata). La “mini-legge sul Corona” ha perso la propria efficacia nell’agosto 2020. Detta legge prevedeva che qualsiasi dichiarazione di emergenza da Covid-19 poteva rimanere in vigore per un massimo di 45 giorni. La stessa, tuttavia, poteva essere prorogata, previa approvazione parlamentare, per periodi di ulteriori 60 giorni. Il Parlamento in qualsiasi momento poteva anche revocare la dichiarazione di emergenza.

L’entrata in vigore della Legge sul Corona ha fatto scattare automaticamente la dichiarazione di emergenza, che è durata 45 giorni, poi prorogata alcune volte. La legge conferiva al Governo, durante lo stato di emergenza, il potere di emanare regolamenti temporanei, se necessario, per prevenire la diffusione del virus. Il Governo doveva considerare le possibili alternative cui poter ricorrere e il loro impatto sui diritti civili e sull’economia. A seconda del tipo di regolamento, le norme derogatorie potevano rimanere in vigore per un massimo di 7, 14 o 28 giorni, dopodiché sarebbero decadute, a meno che non fossero state prorogate dallo stesso Governo. I regolamenti dovevano essere presentati ad una apposita Commissione della Knesset, che poteva approvarli o respingerli entro 24 ore. Se la Commissione non avesse preso provvedimenti entro 24 ore, i regolamenti sarebbero entrati in vigore. Successivamente, la Commissione avrebbe potuto ancora esaminarli, approvandoli o respingendoli entro 14 giorni.

Nel settembre 2020, la Knesset ha approvato l’emendamento n. 2 alla Legge sul Corona. L’emendamento consentiva al Governo di dichiarare: “Una situazione di emergenza speciale a causa del Coronavirus”, nell’ambito di un’emergenza (non speciale) già vigente nel contesto della pandemia da Coronavirus. Il principale cambiamento introdotto da tale emendamento è stata la possibilità, nella “situazione di emergenza speciale”, di imporre restrizioni alle manifestazioni. In uno sviluppo considerato da alcuni come motivato da ragioni politiche piuttosto che sanitarie, l’emendamento n. 2 consentiva al Governo di limitare la partecipazione ad una manifestazione a un chilometro dal proprio domicilio durante una emergenza “speciale” da coronavirus. Al momento dell’entrata in vigore, l’emendamento prevedeva una “emergenza speciale” di sette giorni, prorogata una volta per sondaggio telefonico dai membri del Governo e scaduta il 13 ottobre 2020. La Corte Suprema israeliana ha ritenuto costituzionalmente legittimi sia la Legge sul Corona sia l’emendamento n. 2⁴².

Fino alla Legge sul Corona, i regolamenti emergenziali trovavano apposita copertura costituzionale all’art. 39 della Legge fondamentale israeliana. Attraverso tali regolamenti veniva attuata un’ampia serie di misure, tra cui, ad esempio: autorizzare la sorveglianza e il tracciamento dei contatti, limitare gli spostamenti, gli assembramenti e l’apertura di attività commerciali, imporre chiusure di locali e dichiarare le aree come “zone rosse” (istaurando cordoni sanitari), imporre il primo lockdown nazionale di Israele, imporre la quarantena nelle strutture statali, regolare il funzionamento del

⁴² HCJ, *National Responsibility - Israel My Home Ltd. (PBC) v Government of Israel* [2020] 5469/20 HCJ.

sistema giudiziario durante la pandemia e autorizzare i servizi di sicurezza israeliani ad utilizzare la tecnologia di sorveglianza per la tracciabilità dei contagiati e trasferire le relative informazioni al Ministero della Salute. Tramite i regolamenti emergenziali del Governo alcune violazioni delle ordinanze di sanità pubblica sono state rese illeciti penali o amministrativi. Ai sensi dell'art. 39(6) della Legge fondamentale, i regolamenti emergenziali rimangono in vigore per tre mesi, a meno che non vengano prorogati dalla legge.

Nel gennaio 2022, la Knesset ha approvato una nuova "Legge sul coronavirus", ponendo fine allo stato di emergenza, che è stato in vigore per la maggior parte della pandemia, richiedendo un maggiore controllo legislativo sulle restrizioni legate al COVID-19 imposte dal Governo. La nuova legge è entrata in vigore il 1° febbraio 2022 ed è durata fino alla fine del 2022. In tale nuovo quadro legislativo il Governo per imporre restrizioni relative al coronavirus era tenuto a dichiarare una delle seguenti due situazioni: stato di emergenza o "situazione sanitaria speciale" (caso meno grave). La Knesset doveva approvare le decisioni assunte dal Governo in entrambe le ipotesi. In base alla nuova legge, la Commissione per la Costituzione, la Legge e la Giustizia della Knesset o l'intero plenum dovevano approvare una delle due condizioni sanitarie nazionali entro sette giorni. Se nessuno dei due avesse proceduto in questo senso, il provvedimento adottato dal Governo sarebbe decaduto.

La legge consentiva al Governo di dichiarare una situazione sanitaria speciale solo se i ministri erano convinti che un'ondata attuale del virus avrebbe messo a rischio la salute pubblica. Per dichiarare uno stato di emergenza, i ministri, invece, dovevano essere persuasi che sussistesse una tendenza significativa al rialzo nella morbilità o che il virus fosse abbastanza diffuso da rischiare di destabilizzare l'intero sistema sanitario pubblico. In entrambi i casi, tuttavia, il Governo doveva ricevere i pareri del Ministro della Salute e del personale professionale del medesimo Ministero. Ogni stato di emergenza imposto poteva avere una durata massima di 45 giorni, mentre le situazioni sanitarie speciali potevano durare al massimo 90 giorni. Con la dichiarazione dello stato di emergenza il Governo poteva imporre un lockdown, richiedere che le scuole svolgessero le lezioni online e poteva limitare i viaggi all'estero. Sia nello stato di emergenza che in una situazione sanitaria speciale, il diritto di protestare rimaneva intatto, diversamente da quanto previsto con la legge precedente, che consentiva limiti su tali riunioni durante determinati lockdown.

Per quanto riguarda la situazione sanitaria speciale, le restrizioni potevano entrare in vigore solo dopo che la Commissione per la Costituzione della Knesset avesse votato favorevolmente per la loro adozione. Le norme disciplinanti le restrizioni dovevano essere presentate alla Commissione almeno cinque giorni prima della loro discussione. Nello stato di emergenza, invece, nuove restrizioni potevano essere istituite prima che la Commissione per la Costituzione le approvasse, ma la Commissione aveva 48 ore per ratificare la decisione — 24 ore in più rispetto a quelle previste dalla legge precedente.

Per quanto riguarda la Knesset, dopo una prima fase di assestamento a seguito delle elezioni, ha istituito nuove commissioni parlamentari spe-

ciali, come la Commissione speciale per la gestione del coronavirus⁴³. Le commissioni erano presiedute da membri dei partiti di opposizione. La Knesset ha anche adattato le proprie procedure per consentire un efficace funzionamento del Parlamento, nonostante i rischi per la salute e le restrizioni sul distanziamento sociale. Il plenum e le commissioni hanno continuato a riunirsi fisicamente con regolarità. Le commissioni sono state, inoltre, relativamente rigorose nell'esercitare il loro ruolo di supervisione, nel periodo che va dalla fine di marzo alla metà di maggio 2020. Tuttavia, verso la fine di luglio 2020, con l'adozione della Legge sul Corona, sono stati conferiti ampi poteri al Governo e il controllo parlamentare sul suo operato si è affievolito a fronte di una nuova maggioranza filogovernativa. Il controllo giurisdizionale da parte della Corte Suprema non è stato limitato durante la pandemia⁴⁴. Questo è un punto fondamentale in quanto la Corte Suprema israeliana opera anche in qualità di Alta Corte di Giustizia, esercitando un controllo di diritto amministrativo e costituzionale sull'operato degli altri poteri dello Stato. Nel corso del 2020, dopo una fase iniziale in cui la Corte Suprema si è limitata ad assicurare profili più procedurali del diritto costituzionale dell'emergenza, dall'inizio del 2021, sembra esserci stato un certo cambiamento nell'approccio dell'organo di giustizia costituzionale israeliano, con una sua maggiore propensione ad invalidare le misure relative al Covid-19, su basi normative sostanziali⁴⁵. Nel febbraio 2021, per la prima volta, la Corte ha dichiarato incostituzionale uno dei regolamenti del governo sul Covid-19⁴⁶, e da allora sono seguiti diversi ulteriori interventi sulle misure adottate dal Governo⁴⁷. Un esempio emblematico di questo mutamento nell'approccio della Corte è proprio la decisione riguardante le misure di sorveglianza tecnologica del Governo, che autorizzavano il GSS (Servizi di sicurezza nazionale) a tracciare la posizione dei cittadini israeliani contagiati con strumenti tecnologici. In linea con la periodizzazione sopra ricordata, nella sua prima decisione in merito, datata aprile 2020, la Corte ha adottato un approccio più procedurale e maggiormente basato sul principio di separazione dei poteri. Ha così stabilito che una misura così vasta (attribuzione ai Servizi di sicurezza nazionale della competenza al tracciamento dei contagi) può essere autorizzata solo dalla Knesset attraverso una fonte normativa di rango primario. Ha aggiunto anche indicazioni su adeguate garanzie procedurali, come l'inclusione di clausole di decadenza (cc.dd. *sunset clauses*) nella legislazione.

⁴³ O. Akirav - I. Bar-Siman-Tov, *The Knesset in the Pandemic: A Parliament Amid Weak Institutional Settings and Overlapping Crises*, in Bar Ilan University Faculty of Law Research Paper No. 4891017, in *ssrn.com*.

⁴⁴ Y. Dotan, *The Israeli High Court of Justice during the COVID-19 Crisis: The Model of Continuous Judicial Review*, in *Georgia Law Review*, 58, 2024, 1213 ss.

⁴⁵ I. Bar-Siman-Tov - C. Itay - K. Chani, *The Changing Role of Judicial Review During Prolonged Emergencies: The Israeli Supreme Court During COVID-19*, in *Legal Policy and Pandemics: The Journal of the Global Pandemic Network*, 1(1-3), 2021, 271 ss.

⁴⁶ HCJ, *Idan Mercatz Dimona Ltd.v. Government of Israel*, HCJ 6939/20, 2 February 2021.

⁴⁷ HCJ, *Association for Civil Rights in Israel v Knesset*, HCJ 6732/20; *Oren Shemesh v Prime Minister* HCJ 1107/21 (HCJ). Vedi anche, *Physicians for Human Rights v Minister of Public Security* HCJ 158/21 (HCJ), in cui la Corte ha stabilito che la decisione del Ministro della Sicurezza Pubblica di proibire al Servizio Penitenziario israeliano di vaccinare i detenuti era incostituzionale.

Nella sua seconda decisione, del marzo 2021, la Corte ha effettuato un controllo costituzionale maggiormente sostanziale delle decisioni adottate dal Governo sulla base della legge emanata dal Parlamento per autorizzare la sorveglianza del GSS, e ha stabilito che l'uso estensivo della sorveglianza del GSS violava il diritto alla privacy dei consociati.

4.2.2. *Contact tracing* e tutela dei principi fondamentali nell'esperienza israeliana

Israele ha scelto un approccio peculiare rispetto a quello adottato da altri Paesi, utilizzando il proprio servizio di sicurezza nazionale per garantire il tracciamento dei contatti, una modalità considerata estrema rispetto a quelle cui sono ricorse altre nazioni democratiche⁴⁸. Lo *Shin Bet* è infatti l'agenzia di sicurezza interna israeliana responsabile della protezione di Israele dal terrorismo. Fin dall'inizio della pandemia, lo *Shin Bet* è stato coinvolto nell'ambito delle politiche sanitarie sviluppate dal Governo. La crisi sanitaria è stata trattata come una minaccia per la sicurezza pubblica interna, andando così a rafforzare l'idea che la protezione della salute rientri nell'ambito della tutela dell'ordine pubblico. Senza passare dal Parlamento, il governo israeliano ha attribuito allo *Shin Bet* il potere di rintracciare i cittadini attraverso i dati di localizzazione dei telefoni cellulari delle persone risultate positive al Covid o sospettate di esserlo, senza il loro consenso.

Adita la Corte Suprema israeliana per il vaglio di legittimità costituzionale della normativa adottata dal Governo, la stessa si è pronunciata con la sentenza *Ben Meir* del 26 aprile 2020⁴⁹.

Nella propria decisione, la Corte Suprema ha ritenuto che l'uso di tecnologie non trasparenti per monitorare i cittadini costituisse una grave violazione del diritto alla privacy. La Corte ha analizzato la situazione nella prospettiva peculiare dettata dall'emergenza sanitaria e ha stabilito che mentre le misure emergenziali potevano essere giustificate nelle prime settimane di pandemia, la situazione doveva essere successivamente riportata alla "regola". L'attribuzione allo *Shin Bet* di competenze che vanno oltre la sicurezza nazionale avrebbe dovuto essere oggetto di una legge specifica adottata dal Parlamento.

Gli argomenti utilizzati dal Governo israeliano per giustificare il ricorso ai servizi di sicurezza si basavano in particolare sull'urgenza delle operazioni per evitare una degenerazione della situazione sanitaria del Paese, ma anche sulla peculiarità che caratterizza la comunità ultraortodossa, per la quale il ricorso alla diversa opzione consistente nell'uso di un'applicazione per il tracciamento sul telefono non avrebbe funzionato, essendo mezzo

⁴⁸ R. Cohen-Almagor - E. Haber, *Tracking and Tracing in Israel during COVID-19*, in *Israel Studies Review*, 38(3), 2023, 29 ss.

⁴⁹ HCJ, *Ben Meir v. Prime Minister*, HCJ 2109/20, 26 April 2020. Sul punto si vedano, ad esempio: T.H. Brandes, *Tracking Citizens: Military Surveillance Tools in Israel and Privacy in a Pandemic*, in *VerfBlog*, 22 marzo 2021; E. Bottini, *Contact tracing et constitutions: les juridictions constitutionnelles face aux défis numériques de la pandémie*, in *Cahiers de la recherche sur les droits fondamentaux*, 21, 2023, 53 ss.

milione le persone che non possiedono uno smartphone. Il governo ha quindi deciso di affidarsi alle capacità tecnologiche dello *Shin Bet* e alla raccolta centralizzata dei dati, obbligatoria per tutti i cittadini. Secondo il Governo, l'efficacia di questa misura avrebbe compensato l'inevitabile violazione della privacy.

Quando una persona risultava positiva al test, il Ministero della Salute forniva tutti i relativi dati identificativi (nome, numero di previdenza sociale, numero di telefono), così consentendo di individuare tutte le persone con cui la stessa era entrata in contatto per almeno quindici minuti nel raggio di due metri nelle due settimane precedenti. Per fare ciò, lo *Shin Bet* si è basato su un database chiamato "*The Tool*", che esiste da tempo in Israele per scopi di pubblica sicurezza e antiterrorismo ed è stato utilizzato in particolare come mezzo per rintracciare i militanti palestinesi ricercati. Tale database raccoglie i dati del traffico telefonico di chiunque utilizzi i servizi di telecomunicazione in Israele. Basandosi sulle informazioni raccolte dallo *Shin Bet*, il Ministero della Salute provvedeva poi ad inviare apposito messaggio alle persone identificate, informandole che dovevano isolarsi e registrarsi nel database del Ministero stesso. Sennonché, tutte queste operazioni venivano svolte senza alcun controllo esterno, senza alcuna convalida giudiziaria e, ovviamente, senza il consenso degli interessati. Solo il Primo Ministro e il Procuratore Generale erano informati sull'andamento delle operazioni di tracciamento, attraverso rapporti classificati come segreti di Stato. In particolare, tutte le norme che regolavano la raccolta dei dati da parte dell'Agenzia per la sicurezza interna erano classificate come "segrete" e l'esistenza stessa del database è rimasta segreta fino all'esplosione della pandemia da Covid-19. Sebbene il Ministero della Salute abbia sviluppato in un secondo momento un'applicazione decentralizzata da installare su base volontaria, questa non è stata ritenuta sufficientemente efficace agli occhi del Governo, al punto da sostituire "*The Tool*".

Per quanto più precisamente riguarda l'intervento della Corte Suprema, inizialmente, nel 2020, la stessa aveva insistito sul rispetto della procedura legislativa per regolarizzare le misure sanitarie, adottando un approccio formalistico basato sulla separazione dei poteri. A partire dal 2021, tuttavia, la Corte ha dimostrato una maggiore inclinazione a censurare le misure impugnate a fronte di violazioni di tipo sostanziale dei diritti fondamentali. Nel 2021, infatti, la Corte Suprema ha emanato una sentenza, con la quale vietava il tracciamento dei telefoni cellulari⁵⁰. Nella propria decisione, la Corte si è detta preoccupata che il tracciamento digitale, imposto come misura temporanea di emergenza, diventasse gradualmente permanente. La Corte ha pertanto fissato limiti temporali volti a porre fine ad un uso indiscriminato della sorveglianza digitale, limitandola ai contagiati che si rifiutavano di cooperare con le indagini epidemiologiche.

A seguito della sentenza della Corte Suprema che stabiliva che, dopo il periodo iniziale di emergenza, il Governo non poteva più autorizzare il GSS ad eseguire il tracciamento dei contagiati, la Knesset ha approvato una norma che autorizzava il Governo a consentire ai servizi di *intelligence* di utilizzare la tecnologia per il tracciamento dei contatti dei cittadini isra-

⁵⁰ HCJ, *Association for Civil Rights in Israel v Knesset*, HCJ 6732/20, 1 March 2021.

eliani a partire dal 1° luglio 2020. Da allora, il Governo ha regolarmente esteso tale attribuzione di competenza al GSS. Il periodo massimo per ogni atto di autorizzazione in ordine al tracciamento era di 21 giorni. Le indagini epidemiologiche sono state inizialmente condotte dal Ministero della Salute e successivamente trasferite all'esercito.

Il caso israeliano mette, dunque, in luce il rischio di una torsione securitaria della gestione emergenziale: la salute viene assimilata alla sicurezza nazionale e le tecnologie di tracciamento diventano strumenti di *intelligence*. Sebbene la Corte Suprema abbia progressivamente riequilibrato il sistema, il dato rimane quello di una forte compressione dei diritti fondamentali, difficilmente compatibile con i principi classici del costituzionalismo

5. Alcune considerazioni sull'uso delle tecnologie emergenti da parte dei poteri pubblici durante gli stati di emergenza alla luce della comparazione delle esperienze belga e israeliana

L'analisi comparata delle esperienze belga e israeliana durante la pandemia da Covid-19 evidenzia due modelli profondamente differenti nell'uso delle tecnologie di tracciamento dei contagi, entrambi tuttavia segnati da tensioni con i principi costituzionali di trasparenza, accessibilità e privacy. In Belgio, il sistema si è strutturato all'interno di un quadro istituzionale multilivello, nel quale il Governo federale e le entità federate hanno cercato di cooperare mediante accordi formali. La Corte costituzionale belga ha svolto un ruolo cruciale nell'individuare le falle del sistema, censurando in particolare l'assenza di limiti temporali alla conservazione dei dati, la scarsa chiarezza sulla responsabilità del trattamento e l'indebita attribuzione di poteri a organi non sufficientemente soggetti a controllo parlamentare. Si è trattato, dunque, di un modello formalmente ancorato al diritto europeo e rispettoso della cornice dello Stato di diritto, ma che ha mostrato quanto sia fragile la garanzia dei diritti fondamentali quando il principio di trasparenza cede di fronte a logiche di urgenza e di efficienza.

Israele, al contrario, ha seguito un approccio di natura eminentemente securitaria, affidando lo strumento del *contact tracing* ai servizi di *intelligence*. In questo caso, l'urgenza della tutela della salute pubblica è stata immediatamente assimilata alla difesa della sicurezza nazionale, con conseguente erosione della privacy e opacità nell'utilizzo delle tecnologie. La Corte Suprema israeliana, inizialmente cauta e concentrata sugli aspetti procedurali, ha progressivamente assunto un ruolo più incisivo, arrivando a dichiarare l'incostituzionalità di un uso esteso e non proporzionato della sorveglianza digitale. Il caso israeliano dimostra come la carenza di trasparenza e di garanzie di accessibilità democratica conduca a un vero e proprio slittamento di paradigma, in cui la gestione della pandemia viene assimilata a un'emergenza di sicurezza interna, con una forte compressione dei diritti individuali.

Il raffronto tra i due modelli mette in luce che, sebbene Belgio e Israele abbiano fatto ricorso a strumenti giuridici e istituzionali molto diversi, entrambi hanno sperimentato forme di erosione della privacy e difficoltà nel

garantire la piena *accountability* dei processi decisionali. La differenza sta nel grado di istituzionalizzazione e controllo: in Belgio il circuito democratico e giudiziario ha mantenuto una funzione correttiva, mentre in Israele la logica emergenziale e securitaria ha prevalso più a lungo, spostando il baricentro delle decisioni fuori dal Parlamento e a favore degli apparati di *intelligence*.

Dalla comparazione delle due esperienze sopra citate emerge che l'uso delle nuove tecnologie durante gli stati di emergenza rappresenta un banco di prova per la tenuta dei principi dello Stato costituzionale. In entrambi i casi analizzati, la pandemia ha spinto i governi a privilegiare soluzioni rapide e tecnologicamente sofisticate, ma la velocità della risposta ha comportato un inevitabile deficit di trasparenza, di accessibilità alle informazioni e di tutela effettiva della privacy. La lezione che se ne trae è che non basta affidarsi alla presunta neutralità della tecnologia: occorre costruire un solido quadro di garanzie, che impedisca alle emergenze di trasformarsi in spazi di eccezione permanente.

Un ulteriore elemento di riflessione critica emerge dall'analisi comparata dei due ordinamenti: l'asimmetria tra strumenti tecnologici globali e cornici giuridiche nazionali. Le piattaforme di *contact tracing* e gli algoritmi utilizzati durante la pandemia operano secondo logiche transnazionali, mentre i meccanismi di garanzia restano ancorati a ordinamenti giuridici statali. Questo squilibrio produce un effetto paradossale: Stati formalmente vincolati al rispetto dei principi costituzionali si trovano a dover disciplinare strumenti che sfuggono, per natura, alla loro piena giurisdizione. In Belgio, tale tensione è stata in parte mediata dal diritto europeo, che funge da cornice comune di protezione dei dati. In Israele, invece, la mancanza di un ancoraggio sovranazionale ha accentuato il carattere securitario delle misure, lasciando alla Corte Suprema l'onere di fungere da contrappeso.

La comparazione suggerisce inoltre che non esiste un modello uniforme di bilanciamento tra tecnologia ed emergenza. La differenza tra un approccio multilivello (Belgio) e uno fortemente centralizzato (Israele) riflette non solo tradizioni giuridiche diverse, ma anche culture costituzionali differenti in ordine alla percezione della sicurezza. Ciò conferma come l'uso delle tecnologie emergenti non sia mai neutro, ma incorpori visioni politiche e costituzionali peculiari, capaci di incidere in modo duraturo sul rapporto tra cittadini e istituzioni.

A questo punto possono essere formulate tre riflessioni generali alla luce della ricostruzione sopra operata. In primo luogo, ogni sistema di tracciamento deve essere accompagnato da regole chiare sulla conservazione e sull'uso dei dati, integrate da meccanismi di controllo giurisdizionale e parlamentare effettivi. Senza un limite temporale e sostanziale, la logica emergenziale rischia di normalizzare restrizioni sproporzionate dei diritti. In secondo luogo, l'esperienza israeliana dimostra che l'affidamento delle tecnologie di tracciamento a organi securitari conduce a una torsione dell'ordinamento: la sanità viene assimilata alla sicurezza nazionale e la privacy diviene un valore negoziabile. È invece essenziale mantenere il circuito democratico come luogo di legittimazione e supervisione delle scelte tecnologiche, evitando che esse siano delegate ad apparati opachi. In terzo luogo, la comparazione conferma che la resilienza del costituzionalismo

dipende dalla capacità di rendere intelligibili le decisioni pubbliche: l'accessibilità e la comprensibilità dei processi sono condizioni di fiducia civica e, dunque, di efficacia delle misure stesse.

Il futuro dello Stato costituzionale di fronte alle emergenze digitali non può quindi ridursi a una contrapposizione tra efficienza e garanzie. La vera sfida è conciliare rapidità e legalità, innovazione tecnologica e trasparenza democratica. Solo così le nuove tecnologie potranno essere strumenti al servizio della collettività, e non leve di erosione dei principi fondamentali. In questo senso, Belgio e Israele, pur con esiti diversi, ci mostrano due facce dello stesso problema: l'urgenza di dotare lo Stato di strumenti capaci di resistere alla tentazione di sospendere, anche solo implicitamente, la grammatica costituzionale di fronte alle crisi. La lezione che ne deriva è che la resilienza costituzionale non consiste nell'evitare la tecnologia, ma nel disciplinarla attraverso limiti chiari, *sunset clauses*, *accountability* parlamentare e controllo giurisdizionale. Solo così le società democratiche potranno evitare che l'emergenza diventi un paradigma ordinario di governo.

Abstract

Il contributo analizza l'uso delle tecnologie emergenti da parte dei governi durante gli stati di emergenza, con particolare riferimento ai casi del Belgio e di Israele nella gestione della pandemia da Covid-19. Se da un lato strumenti digitali come intelligenza artificiale e sistemi di *contact tracing* hanno garantito efficienza e rapidità nelle risposte pubbliche, dall'altro hanno sollevato questioni critiche in merito a trasparenza, responsabilità democratica e tutela dei diritti fondamentali. Il modello belga, pur fondato sulla cooperazione multilivello e su controlli giurisdizionali, ha mostrato fragilità nel garantire limiti certi all'uso dei dati. Diversamente, Israele ha seguito un approccio di matrice securitaria, affidando il tracciamento ai servizi di *intelligence*, con gravi ricadute sulla privacy, successivamente corrette dall'intervento della Corte Suprema. Attraverso la comparazione, lo studio evidenzia come l'adozione delle nuove tecnologie in contesti emergenziali richieda solide garanzie normative e istituzionali per preservare i principi dello Stato costituzionale.

The paper investigates the use of emerging technologies by governments during states of emergency, focusing on the Belgian and Israeli experiences during the Covid-19 pandemic. It highlights how digital tools such as artificial intelligence and contact tracing systems, while enhancing efficiency and rapidity of governmental responses, also pose serious challenges to constitutional principles, especially transparency, accountability, and the protection of fundamental rights. The Belgian model, anchored to European legal frameworks, relied on multilevel cooperation and was subject to judicial corrections by the Constitutional Court. In contrast, Israel adopted a security-driven approach, delegating contact tracing to intelligence services, which resulted in a significant erosion of privacy until the Supreme Court intervened. Through comparative analysis, the study underscores how the integration of technology in emergency governance risks weakening democratic guarantees if not supported by clear legal frameworks and effective judicial and parliamentary oversight.

Keywords

tecnologie emergenti - sorveglianza digitale - Stati di emergenza - resilienza costituzionale - comparazione giuridica

Algorithmic Discrimination in States of Emergency: The Role of EU Law in Promoting Counter-Strategies*

Gabriele Rugani

Table of Contents

1. Introduction. – 2. Algorithmic discrimination in general. – 3. Algorithmic discrimination in states of emergency. – 3.1. Algorithmic discrimination in states of emergency: the case of mass surveillance and registration measures. – 3.2. Algorithmic discrimination in states of emergency: the case of decision-support tools. – 4. The EU legal framework on artificial intelligence under the AI Act. – 5. The strategies promoted by the AI Act to counter algorithmic discrimination, with particular reference to states of emergency. – 5.1. Prohibited AI systems. – 5.2. High-risk AI systems. – 5.3. The fundamental rights impact assessment for high-risk AI systems. – 6. Concluding remarks.

1. Introduction

The present contribution focuses on Regulation (EU) 2024/1689 of the European Parliament and of the Council «laying down harmonised rules on artificial intelligence» and amend-

ing certain EU legislative acts – better known as «Artificial Intelligence Act» or «AI Act» – adopted by EU legislators on 13 June 2024.

The aim of the paper, however, is not to provide an exhaustive analysis of the entire Regulation, but rather to address a much more specific research question: whether, and to what extent, the AI Act attempts to promote strategies to counter algorithmic discrimination, with particular reference to that occurring in states of emergency.

To this end, it will preliminarily be necessary to illustrate the issue of algorithmic discrimination in general: the most affected categories, the reasons

*L'articolo è stato sottoposto, in conformità al regolamento della Rivista, a referaggio “a doppio cieco”.

Il presente contributo è realizzato nell'ambito del PRIN 2020 “Decision-Making in the Age of Emergencies. New Paradigms in Recognition and Protection of Rights [DeMa]” – CUP C61B22002230001.

This article is the result of the research carried out at the University of Pisa (UniPi), in the framework of the Project “PRIN AiGeDi – Artificial Intelligence between generating and tackling gender-based discriminations”.

why AI systems can produce discriminatory effects, and the consequences of such discrimination will be identified in brief.

Subsequently, the focus will shift to the matter of algorithmic discrimination in states of emergency: in particular, the reasons why this phenomenon raises specific concerns and deserves particular attention will be explained. For this purpose, the paper will take into consideration two of the main categories of digitalised and AI-based solutions most frequently employed during crises: on the one hand, mass surveillance and registration measures; on the other, decision-support tools. It will be explored whether and to what degree such solutions can produce discriminatory effects.

At this point, it will be possible to turn to the AI Act. First, the essential features of this instrument will be outlined. Next, the contribution will examine the presence of provisions aimed at countering not only algorithmic discrimination in general, but also that which can occur in states of emergency.

Finally, the contribution will offer some concluding remarks, highlighting both the reasons why the AI Act represents a step forward in addressing algorithmic discrimination, including in emergencies, and the limitations and margins for improvement that the EU legal framework still presents in this regard.

2. Algorithmic discrimination in general

It has long been widely demonstrated – through statistical analyses (such as those by the World Economic Forum¹ and UNESCO²), academic studies (including the «Gender Shades» project on facial recognition technologies³) and well-known cases (such as the algorithm developed by Amazon for recruitment purposes⁴) – that AI can produce discriminatory effects against women, as well as various minority and vulnerable groups⁵. This includes, for example, individuals belonging to racial or ethnic minorities, persons with disabilities, and people who are not heterosexual or cisgender⁶.

¹ See for instance: World Economic Forum, *The Global Gender Gap Report*, Cologne/ Geneva, 2018.

² See for instance: United Nations Educational, Scientific and Cultural Organization (UNESCO), *Artificial Intelligence and Gender Equality*, Paris, 2020.

³ J. Buolamwini – T. Gebru, *Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification*, in *Proceedings of Machine Learning Research*, 81, 2018, 77 ff.; J. Buolamwini, *Hearing on: Artificial Intelligence: Societal and Ethical Implications – United States House Committee on Science, Space and Technology*, Washington DC, 2019.

⁴ On the point see, for instance: J. Dustin, *Amazon scraps secret AI recruiting tool that showed bias against women*, in *Reuters*, 11 October 2018; J. Lauret, *Amazon's sexist AI recruiting tool: how did it go so wrong?*, in *becominghuman.ai*, 16 August 2019.

⁵ On the point see, for instance: C. Nardocci, *Intelligenza artificiale e discriminazioni*, in *Rivista "Gruppo di Pisa"*, 3, 2021, 9 ff.

⁶ According to the Cambridge Dictionary, the term «cisgender» is «used to describe a person whose gender matches the body they were born with». Consequently, a non-

Individuals who face multiple factors of discrimination simultaneously – e.g., a woman belonging to an ethnic-racial minority – are particularly disadvantaged. As is now well known, this phenomenon is referred to as «intersectional discrimination»⁷: this concept, originating in North America⁸, has subsequently been taken into consideration at the United Nations level⁹, as well as by European Union institutions. In particular, with its Resolution of 6 July 2022 dedicated specifically to «intersectional discrimination», the European Parliament defined it as «a situation in which several grounds of discrimination operate and interact with each other, for example gender with other grounds of discrimination, such as race, colour, ethnic or socio-economic status, age, sexual orientation, gender identity and expression, sex characteristics, genetic features, religion or belief, nationality, residence status, migrant background, or disability, among others, in a way that is inseparable and produces specific types of discrimination»¹⁰.

Returning to AI, the reasons why it can produce the aforementioned discriminatory effects may lie in several factors¹¹. Among these, the most

cisgender person is someone for whom such correspondence does not occur.

⁷ On the topic see, for instance: C.A. MacKinnon, *Intersectionality as method: A note*, in *Signs: Journal of Women in Culture and Society*, 38(4), 2013, 1019 ff.; H. Lutz, *Intersectionality as method*, in *DiGeSt. Journal of Diversity and Gender Studies*, 2(1)-2, 2015, 39 ff.; J. Bailey – V. Steeves – J. Burkell – L.R. Shade – R. Ruparelia – P. Regan, *Getting at equality: Research methods informed by the lessons of intersectionality*, in *International Journal of Qualitative Methods*, 18, 2019, 1 ff.; S. Grabe, *Research methods in the study of intersectionality in psychology: Examples informed by a decade of collaborative work with majority World Women's Grassroots Activism*, in *Frontiers in Psychology*, 11, 2020; C. Haynes – N.M. Joseph – L.D. Patton – S. Stewart – E.L. Allen, *Toward an understanding of intersectionality methodology: A 30-year literature synthesis of black women's experiences in higher education*, in *Review of Educational Research*, 90(6), 2020, 751 ff.; A.J. Petrozziello, *Intersectionality as method for human rights research: Identifying who is made stateless and how through UN treaty body reviews*, in *Journal of Human Rights*, 24(2), 2025, 182 ff.

⁸ The idea of «intersectionality» was first coined in: K. Crenshaw, *Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics*, in *University of Chicago Legal Forum*, 1, 1989, 139 ff.

⁹ United Nations Committee on the Elimination of Discrimination against Women (CEDAW), *General recommendation No. 28 on the core obligations of States parties under article 2 of the Convention on the Elimination of All Forms of Discrimination against Women*, CEDAW/C/GC/28, 16 December 2010, para. 18, according to which: «Intersectionality is a basic concept for understanding the scope of the general obligations of States parties [...]. The discrimination of women based on sex and gender is inextricably linked with other factors that affect women, such as race, ethnicity, religion or belief, health, status, age, class, caste and sexual orientation and gender identity. Discrimination on the basis of sex or gender may affect women belonging to such groups to a different degree or in different ways to men. States parties must legally recognize such intersecting forms of discrimination and their compounded negative impact on the women concerned and prohibit them. They also need to adopt and pursue policies and programmes designed to eliminate such occurrences, including, where appropriate, temporary special measures [...]».

¹⁰ European Parliament, *Resolution on intersectional discrimination in the European Union: the socio-economic situation of women of African, Middle-Eastern, Latin-American and Asian descent (2021/2243(INI))*, 6 July 2022.

¹¹ On the point see, for instance: S. Barocas – A.D. Selbst, *Big data disparate impact*, in *California Law Review*, 104, 2016, 671 ff.; F.Z. Borgesius, *Discrimination, artificial intelligence, and algorithmic decision-making*, Strasbourg, 2018, 10 ff.; P. Zuddas, *Intelligenza*

significant is undoubtedly the so-called «data training», i.e. the collection and selection of data, since the more biased or prejudiced the data are, the more the model will reflect such bias, acting in a discriminatory manner: this is often described as the «garbage in, garbage out» rule¹². Other relevant factors include the following: the definition of, and the relationships established between, the characteristic the system looks for («target variable») and the category associated with it («class label»)¹³; the selection of individual features which are relevant for building the model («feature selection»)¹⁴; the choice of the so-called «proxy», i.e. the element based on which the machine differentiates¹⁵; and intentional discrimination, meaning discrimination consciously introduced by the programmer through a deliberate and biased selection of data that has a negative impact on one or more protected groups (a practice referred to as «masking»)¹⁶.

In any case, regardless of the specific causes, algorithms which perpetuate – or even amplify – existing forms of discrimination inevitably contribute to reinforcing social inequalities to the detriment of women, as well as various minority and vulnerable groups, across a wide range of contexts. Among these, we can certainly mention education, employment, health-care and essential services, law enforcement, immigration management and beyond.

artificiale e discriminazioni, in *Consulta Online*, 2020, 1 ff.; C. Nardocci, *Intelligenza artificiale e discriminazioni*, cit., 17 ff.

¹² On the point see, for instance: G. Resta, *Governare l'innovazione tecnologica: decisioni algoritmiche, diritti digitali e principio di uguaglianza*, in *Politica del diritto*, 2, 2019, 199 ff., 214; A. Simoncini, *L'algoritmo incostituzionale: intelligenza artificiale e il futuro delle libertà*, in *BioLaw Journal – Rivista di BioDiritto*, 1, 2019, 63 ff., 84.

¹³ AI techniques operate on the basis of correlations between the two abovementioned parameters: when the first (the «target variable») occurs, the AI system associates it with a specific category (the «class label»), thus making a decision. The issue arises when the system is asked to identify and select characteristics to be automatically linked to categories that are difficult to define, such as a «good employee». The choices made in this regard during the model development phase may be unreasonable and discriminatory, thereby producing harmful effects for protected groups. On the point see, for instance: S. Barocas – A.D. Selbst, *Big data disparate impact*, cit., 677 ff.; T.B. Gillis – J.L. Spiess, *Big Data and Discrimination*, in *The University of Chicago Law Review*, 86(2), 2019, 459 ff., 475.

¹⁴ A partial delimitation that fails to take all relevant aspects into account, or that gives disproportionate weight to some of them at the expense of others, may result in the system functioning in a potentially discriminatory manner.

¹⁵ The choice of the proxy – though it should not fall on traditional grounds of discrimination – may nonetheless be predictive of them: the more the selected proxy reflects characteristics commonly shared by members of a protected group, the more likely it is that the system, by differentiating on the basis of that proxy, will subject individuals belonging to that group to discriminatory treatment.

¹⁶ On the point see, for instance: F.Z. Borgesius, *Discrimination, artificial intelligence, and algorithmic decision-making*, cit., 13 ff.; J.A. Kroll – J. Huey – S. Barocas – E.W. Felten – J.R. Reidenberg – D.G. Robinson – H. Yu, *Accountable algorithms*, in *University of Pennsylvania Law Review*, 165, 2017, 633 ff.

3. Algorithmic discrimination in states of emergency

Given the above regarding algorithmic discrimination in general, it is now necessary, as anticipated, to investigate how this takes shape in states of emergency.

A trend that has been clearly demonstrated by past emergencies – above all, the CoViD-19 pandemic – is that, in times of crisis, the state-business complex tends to increase its reliance on digitalised and AI-based solutions¹⁷.

Let us go step by step. As is well known, when everyone feels at risk, the acceptability of exceptional measures is higher. It comes as no surprise that extraordinary situations tend to warrant extensions of discretionary governmental powers, which can also be employed as rationales – or pretexts – to suspend and undermine democratic principles and rights¹⁸. This view evokes the theories of the Italian philosopher Giorgio Agamben concerning the state of exception, according to which the emergency has become «the dominant paradigm of government in contemporary politics» via the unusual extension of governmental powers to tackle it¹⁹.

Coming now to what is of even greater relevance for the purposes of the present analysis, it must be specifically highlighted that, among the exceptional measures referred to above, the ones that now predominate are – as anticipated – enhanced digitalised and AI-based solutions, perceived as particularly quick and efficient. It is no coincidence that the business model underpinning the current digital world has benefited from states of emergency by significantly expanding digital giants' profits and their behavioural data supply chains²⁰. This is perfectly consistent with the «surveillance capitalism» theory, first advanced by the US-American scholar Shoshana Zuboff, that describes a new economic order in which human experience – extracted through pervasive digital monitoring – is treated as free raw material to be analysed, commodified, and monetised, primarily for predicting and manipulating human behaviour²¹. Within such

¹⁷ On the point see, for instance: R. Csernatonì, *New states of emergency: normalizing techno-surveillance in the time of COVID-19*, in *Global Affairs*, 2020, 1 ff., 6.

¹⁸ *Ibid.*, 7.

¹⁹ G. Agamben, *Homo Sacer. Sovereign Power and Bare Life*, Stanford, 1998; G. Agamben, *State of Exception*, Chicago, 2005. According to G. Agamben, «the state of exception is neither external nor internal to the juridical order, and the problem of defining it concerns precisely a threshold, or a zone of indifference, where inside and outside do not exclude each other but rather blur with each other» (G. Agamben, *State of Exception*, cit., 23). In brief, it is possible to define the state of exception as a special condition in which, due to an emergency or a serious threat to the state, the juridical order is suspended yet paradoxically remains in force, thereby creating a blurred zone. In such a situation, individuals are reduced to «bare life», exposed to sovereign power without legal protection. Although originally conceived as temporary, the state of exception has become «the dominant paradigm of government in contemporary politics» (G. Agamben, *State of Exception*, cit., 2).

²⁰ R. Csernatonì, *New states of emergency*, cit., 6.

²¹ S. Zuboff, *The Age of Surveillance Capitalism. The fight for a human future and the new frontier of power*, New York, 2019. In greater detail, «surveillance capitalism» is defined as follows:

framework, individuals become the ultimate post-industrial and profitable commodity, and the datafication of their personal information is seen as a legitimate strategy to access, understand, monitor, and predict people's behaviour. Surveillance digital enterprises, such as Google, Amazon, Meta and others, have unparalleled control over knowledge about individuals, groups and society across borders. This introduces massive asymmetries and hierarchies in knowledge that can be leveraged into new forms of power and algorithmic-driven computational governance under the aegis of private capital. It also enables Big Tech companies to play an increasingly geopolitical role²².

Now then, as already said, it has become clear that times of crisis significantly benefit the economic order described above. With reference to such system, it is possible to apply the considerations made by Canadian author Naomi Klein: the latter, indeed, describes what can be defined as «disaster capitalism»²³, the main feature of which is the exploitation of disasters, treated as exciting market opportunities²⁴.

During states of emergency, several distinct trends can be observed with regard to the use of digitalised and AI-based solutions. First, crises often stimulate the development of new *ad hoc* technologies, specifically designed to address the extraordinary circumstances at hand. Second, already existing tools are frequently deployed more extensively. Third, pre-existing technologies are often “repurposed”, that is, converted from their original intended use and redirected toward objectives linked to the emergency. What has just been said will be clarified in greater detail in the following analysis.

Given the trends just outlined, in times of crisis both the likelihood of the algorithmic discrimination discussed earlier and the associated risks are significantly amplified. On the one hand, the sheer increase in the deployment of technologies capable of producing discriminatory outcomes heightens the probability of such effects. On the other hand, during emergencies these technologies are especially applied in highly sensitive contexts – for instance, in the case of the CoViD-19 pandemic, within the healthcare sector – where the consequences of discriminatory decisions

«1. A new economic order that claims human experience as free raw material for hidden commercial practices of extraction, prediction, and sales; 2. A parasitic economic logic in which the production of goods and services is subordinated to a new global architecture of behavioral modification; 3. A rogue mutation of capitalism marked by concentrations of wealth, knowledge, and power unprecedented in human history; 4. The foundational framework of a surveillance economy; 5. As significant a threat to human nature in the twenty-first century as industrial capitalism was to the natural world in the nineteenth and twentieth; 6. The origin of a new instrumentarian power that asserts dominance over society and presents startling challenges to market democracy; 7. A movement that aims to impose a new collective order based on total certainty; 8. An expropriation of critical human rights that is best understood as a coup from above: an overthrow of the people's sovereignty».

²² R. Csernaton, *New states of emergency*, cit., 6.

²³ N. Klein, *The Shock Doctrine: The Rise of Disaster Capitalism*, New York, 2007.

²⁴ *Ibid.*, 19. In greater detail, N. Klein defines «disaster capitalism» as follows: «orchestrated raids on the public sphere in the wake of catastrophic events, combined with the treatment of disasters as exciting market opportunities».

can be particularly severe.

In the following sections of this contribution, two of the main categories of digitalised and AI-based solutions most frequently employed during states of emergency will be considered, without claiming to be exhaustive. The first one encompasses mass surveillance and registration measures, such as contact tracing apps, like those used to contain CoViD-19²⁵. The second one includes decision-support tools, for instance those applied at the clinical level. With regard to both categories, the reasons why algorithmic discrimination may occur, as well as the resulting harmful effects, will be highlighted.

3.1. Algorithmic discrimination in states of emergency: the case of mass surveillance and registration measures

As is well known, mass surveillance and registration measures have often been likened to the English philosopher Jeremy Bentham's prison design, i.e. the so-called «panopticon». This prison design is circular, has a guard tower in its centre and cells along the circumference of the structure. The placement of the tower enables guards to observe every cell in the prison without being seen. The purpose of this prison design is to control inmate behaviour by giving the impression of continuous surveillance. Inmates believe that they are continuously monitored, as they cannot verify the presence or absence of the guards in the central tower. They therefore avoid undesirable behaviour and behave as if they are constantly being watched. Panoptic measures thus monitor all in order to compel conformity and compliance²⁶.

Today, this notion may be complemented by the concepts of «liquid surveillance» and «post-panopticon», which may serve as analytical tools to explore new modes of surveillance that are not determined by top-down and direct state monitoring – as was the case with the original panopticon, as described above – but by indirect observation, algorithmic-driven prediction, and self-surveillance²⁷. In particular, self-surveillance has become a staple feature during the pandemic, reflected by the use of digital technologies such as the cited contact tracing apps by obedient citizens, deputised or compelled by corporate and governmental strategies to perform what can be defined as the «securitisation footwork» of personal health monitoring and self-governance²⁸. This allowed both public and private sector entities to effectively monitor the population during the CoViD-19

²⁵ On CoViD-19 contact tracing apps, see in greater detail: G. Rugani, *Le condizioni ricavabili dal Regolamento generale sulla protezione dei dati per le applicazioni nazionali di tracciamento dei contatti: alcune considerazioni*, in *European Papers*, 5(1), 2020, 633 ff.

²⁶ On Bentham's panopticon, see for instance: J. Semple, *Bentham's Prison: A Study of the Panopticon Penitentiary*, Oxford, 1993.

²⁷ Z. Bauman – D. Lyon, *Liquid surveillance: A conversation*, Cambridge, 2012.

²⁸ P. Mantello, *The machine that ate bad people: The ontopolitics of the precrime assemblage*, in *Big Data & Society*, 2016, 1 ff., 2.

emergency.

Another important concept for our purposes, closely connected to the previous ones, is the notion of «panoptic sort»: the latter describes the process whereby individuals of the population are sorted on the basis of their perceived political, economic, or social worth or value²⁹. Indeed, it can be convincingly argued that the ultimate purpose of surveillance – whether exercised through the traditional forms of the «panopticon» or through the less conventional modalities of the «post-panopticon» – is classification, more specifically «social sorting»³⁰: surveillance technologies are intrinsically designed to identify, assess, classify and sort members of the population into defined categories or groups according to predetermined criteria, with the aim to manage, govern, and influence behaviour based on categorisations³¹. These categories are thus formal mechanisms used to socially control groups and induce their conformity. The categories created because of social sorting and the persons within these categories can shift over time, especially when the boundaries of the categories are redefined³².

The above considerations bring us to the issue of discrimination. Indeed, it must be highlighted that social sorting can discriminate against some individuals, while privileging others³³, and that digital tools allowing for mass data collection and surveillance have played a particularly powerful role in deepening inequalities of discriminatory public policy³⁴. According to the US-American political scientist Virginia Eubanks, «[p]eople of colour, migrants, unpopular religious groups, sexual minorities, the poor, and other oppressed and exploited populations bear a much higher burden of monitoring and tracing than advantaged groups»³⁵.

The fact that marginalised groups are disproportionately targeted by surveillance, which at first glance may seem surprising, is in truth by no means so, and is primarily due to two factors: on the one hand, the fact that their status is perceived by surveillance and registration technologies to be «high-risk»; on the other hand, because their marginality makes resistance to surveillance particularly difficult³⁶.

²⁹ O. H. Gandy Jr., *The Panoptic Sort: A Political Economy of Personal Information*, Boulder/Oxford, 1993.

³⁰ D. Lyon (ed.), *Surveillance as Social Sorting: Privacy, risk, and digital discrimination*, London/New York, 2003.

³¹ D. Lyon, *Why Where You are Matters: Mundane Mobilities, Transparent Technologies, and Digital Discrimination*, in T. Monahan (ed.), *Surveillance and Security: technological Politics and Power in Everyday Life*, New York/London, 2006, 209 ff.

³² M.H. Maras – W. O'Brien, *Discrimination, stigmatization, and surveillance: COVID-19 and social sorting*, in *Information & Communications Technology Law*, 2020, 1 ff., 10.

³³ On the point, see for instance: N. Lewis, *Expanding surveillance: connecting biometric information systems to international police cooperation*, in E. Zureik – M.B. Salter (eds.), *Global Policing and Surveillance: Borders, Security, Identity*, Cullompton, 2005, 97 ff.

³⁴ M.H. Maras – W. O'Brien, *Discrimination, stigmatization, and surveillance*, cit., 11.

³⁵ V. Eubanks, *Automating Inequality: How High-Tech Tools Profile, Police and Punish the Poor*, New York, 2017, 6.

³⁶ On the point see, for instance: J. Gilliom, *Overseers of the poor: Surveillance, resistance, and the limits of privacy*, Chicago, 2001; T. Monahan, *Regulating belonging: Surveillance, inequality,*

Two illustrative examples may be provided. The first concerns an Australian study, published in 2020 by three social scientists³⁷, which shows that homeless people risk becoming «a primary target of data-driven scrutiny»³⁸, *inter alia* because their «need for basic services subverts their capacity to avoid surveillance»³⁹. The second example is an investigation carried out in the United States, according to which individuals from lower-income households – who are often Hispanic and black people – are more reliant on mobile devices as their primary source of internet access than members of wealthier families, who are typically white⁴⁰. However, smartphones are also among the primary tools through which people are tracked: the final result, therefore, is increased surveillance of vulnerable groups.

From surveillance and social sorting processes, therefore, profoundly negative consequences may arise for marginalised individuals, since the categories created through social sorting may lead to overtly differential treatment and adverse material impacts on their life choices and opportunities, including compromised access to education, employment, travel, leisure and other activities and services⁴¹.

During states of emergency – in which mass surveillance and registration measures are significantly increased, as is clear when considering those implemented in response to the CoViD-19 pandemic – the phenomenon described is further expanded and exacerbated⁴². Indeed, such measures are hardly class, race, ethnicity, and gender neutral⁴³. In what Virginia Eubanks describes as «a feedback loop of injustice», marginalised groups are therefore subjected to ever-intensified levels of surveillance, which then reinforces marginality, including by informing automated and embodied decisions about access to essential services⁴⁴.

In conclusion, based on the foregoing, it can be clearly stated that the widespread use of mass surveillance and registration measures, which is typically carried out in response to crisis situations, has the undeniable potential to deepen discrimination against already marginalised groups⁴⁵.

and the cultural production of abjection, in *Journal of Cultural Economy*, 10(2), 2017, 191 ff.

³⁷ A. Clarke – C. Parsell – L.N. Lata, *Surveilling the marginalised: How manual, embodied and territorialised surveillance persists in the age of 'dataveillance'*, in *Sociological Review*, 2020, 1 ff.

³⁸ *Ibid.*, 4.

³⁹ *Ibid.*, 15.

⁴⁰ M. Madden, *Privacy, Security, and Digital Inequality: How Technology Experiences and Resources Vary by Socioeconomic Status, Race, and Ethnicity*, New York, 2017, 8, according to which: «There are also substantial gaps across racial, ethnic, and nativity groups when looking at reliance on mobile connectivity. While 70% of foreign-born Hispanics who own smartphones and use the internet say that most of their internet use takes place on their cell phone, just 34% of white smartphone owners, 52% of blacks, and 53% of U.S.-born Hispanics say this».

⁴¹ M.H. Maras – W. O'Brien, *Discrimination, stigmatization, and surveillance*, cit., 11.

⁴² *Ibid.*

⁴³ *Ibid.*, 16.

⁴⁴ V. Eubanks, *Automating Inequality*, cit., 7.

⁴⁵ M.H. Maras – W. O'Brien, *Discrimination, stigmatization, and surveillance*, cit., 16.

3.2. Algorithmic discrimination in states of emergency: the case of decision-support tools

During states of emergency, among the exceptional measures most frequently relied upon consisting of digitalised and AI-based solutions, it is worth mentioning decision-support tools aimed at enabling quicker and more efficient action. However, their use may also give rise to discriminatory effects against vulnerable people.

Let us begin by considering the healthcare sector. As is well known, AI systems hold promise for improved diagnostic and prognostic decision support, epidemiological monitoring and prediction, and also vaccine discovery⁴⁶. However, much research has reported that these systems can discriminate between, and create unequal outcomes in, different sociodemographic groups⁴⁷.

This is due to several factors. In particular, the use of biased AI models might be disproportionately harmful to vulnerable groups – who are already subject to widespread health inequality – because they are not properly represented in training datasets⁴⁸. Indeed, it must be highlighted that the prevalence and incidence of diseases and their risk factors often vary by population group. If datasets do not adequately cover populations at particular risk, trained prediction models that are used in clinical AI decision support might have lower sensitivity for these populations⁴⁹. Every time a prediction model which has been tailored to the members of a dominant group is applied in a «one-size-fits-all» manner to a disadvantaged group, the model might yield suboptimal results and be harmful for disadvantaged people⁵⁰.

Moreover, during states of emergency, demand for rapid response technological interventions might hinder responsible AI design and use⁵¹. With particular reference to the CoViD-19 pandemic, certain studies have found that owing to the pressure of rushed research, the proposed systems were

⁴⁶ On the point see, for instance: M.N. Islam – T.T. Inan – S. Rafi – S.S. Akter – I.H. Sarker – A.K. Islam, *A survey on the use of AI and ML for fighting the COVID-19 pandemic*, in *arxiv.org*, 2020; H.B. Syeda – M. Syed – K.W. Sexton – S. Syed – S. Begum – F. Syed – F. Prior – F. Yu Jr., *Role of Machine Learning Techniques to Tackle the COVID-19 Crisis: Systematic Review*, in *JMIR Med Inform*, 9(1), 2021.

⁴⁷ On the point see, for instance: M. Ghassemi – T. Naumann – P. Schulam – A.L. Beam – I.Y. Chen – R. Ranganath, *A Review of Challenges and Opportunities in Machine Learning for Health*, in *AMIA Jt Summits Transl Sci Proc*, 2020, 191 ff.

⁴⁸ D. Leslie – A. Mazumder – A. Peppin – M.K. Wolters – A. Hagerty, *Does “AI” stand for augmenting inequality in the era of covid-19 healthcare?*, in *British Medical Journal*, 2021, 1 ff., 1.

⁴⁹ A. Rajkomar – M. Hardt – M.D. Howell – G. Corrado – M.H. Chin, *Ensuring Fairness in Machine Learning to Advance Health Equity*, in *Ann Intern Med*, 169(12), 2018, 866 ff.

⁵⁰ H. Suresh – J.V. Guttag, *A Framework for Understanding Unintended Consequences of Machine Learning*, in *arxiv.org*, 2019.

⁵¹ On the point see, for instance: E. Rösli – B. Rice – T. Hernandez-Boussard, *Bias at warp speed: how AI may contribute to the disparities gap in the time of COVID-19*, in *Journal of the American Medical Informatics Association*, 28(1), 2021, 190 ff.; D. Leslie, *Tackling COVID-19 through responsible AI innovation: five steps in the right direction*, in *Harvard Data Science Review*, 2020.

at high risk of statistical bias, poorly reported, and overoptimistic⁵². In this context, it is worth citing one example in particular. Some hospitals hurriedly repurposed AI systems – which had been developed for use, and trained on data, in situations other than the pandemic – for sensitive tasks such as predicting the deterioration of infected patients who might have required intensive care or mechanical ventilation⁵³. These models ran considerable risks of insufficient validation, inconsistent reliability, and poor generalisability due to unrepresentative samples and a mismatch between the population represented in the training data and those who were disparately affected by the outbreak⁵⁴.

Moreover, once again during the CoViD-19 pandemic, AI systems were swiftly repurposed in non-clinical domains as well, with similarly concerning effects. In an attempt to curb the spread of CoViD-19, the United States prison system, for example, used an algorithmic tool developed for measuring the risk of recidivism to determine which inmates would be released to home confinement. This tool exhibited racial biases, and so repurposing it for the management of health risks made black inmates more likely to remain confined and, consequently, subjected to increased exposure to CoViD-19 infection and disease-related death⁵⁵.

In brief, it is clear that digital and AI-based decision-support tools, often used during crisis situations, while providing valuable contributions in addressing the emergency, can, however, lead to discriminatory decisions, often because vulnerable groups are not adequately represented in the datasets.

4. The EU legal framework on artificial intelligence under the AI Act

After demonstrating not only that AI can produce discriminatory effects, but also that such effects risk being further amplified during states of emergency, it is now possible to focus on how these issues are addressed in the relevant legislation of the European Union, with particular reference to the AI Act.

At this stage, it is appropriate both to briefly retrace the process that led to the adoption of this instrument and to outline its essential content, as this will be of relevance for the subsequent analysis.

On 19 February 2020, in its White Paper «On Artificial Intelligence – A European approach to excellence and trust»⁵⁶, the European Commis-

⁵² L. Wynants – B. Van Calster – G.S. Collins and others, *Prediction models for diagnosis and prognosis of covid-19 infection: systematic review and critical appraisal*, in *British Medical Journal*, 2020, 1 ff.

⁵³ R. Robbins, *‘Human experts will make the call’: Stanford launches an accelerated test of AI to help care for Covid-19 patients*, in *STAT News*, 2020.

⁵⁴ K. Singh – T.S. Valley – S. Tang and others, *Validating a widely implemented deterioration index model among hospitalized COVID-19 patients*, in *medRxiv*, 2020.

⁵⁵ R. Fogliato – A. Xiang – A. Chouldechova, *Why PATTERN should not be used: the perils of using algorithmic risk assessment tools during COVID-19*, in *Partnership on AI*, 2020.

⁵⁶ European Commission, White Paper On Artificial Intelligence – A European

sion highlighted the need for a «regulatory framework» on artificial intelligence, in order to «build trust among consumers and businesses in AI, and therefore speed up the uptake of the technology»⁵⁷. Consequently, after a stakeholder consultation that took place between February and June 2020⁵⁸, on 21 April 2021 the Commission presented the Proposal for a Regulation of the EP and of the Council «laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts»⁵⁹. On 6 December 2022, the Council adopted its «General approach» on the AI Act, while on 14 June 2023 the European Parliament adopted its negotiating position on such instrument⁶⁰. Finally, on 9 December 2023 the two co-legislators reached a political agreement and, on 13 June 2024, Regulation (EU) 2024/1689 of the European Parliament and of the Council «laying down harmonised rules on artificial intelligence» and amending certain EU legislative acts – better known as «Artificial Intelligence Act» – was adopted⁶¹.

With regard to its content, the AI Act – whose «Subject matter» and «Scope» are specified in articles 1 and 2, respectively – sets out the rules applicable to an «AI system», which is defined as «a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments» (art. 3, n. 1). It is important to note, however, that the scope of the Regulation is also subject to certain exceptions: in particular, according to the so-called «exception for military,

approach to excellence and trust, 19 February 2020, COM (2020) 65 final.

⁵⁷ *Ibid.*, § 5.

⁵⁸ On the 2020 stakeholder consultation, see in greater detail: G. Rugani, *Potentialities and Margins for Improvement of the European AI Alliance, an Example of Participatory Democracy in the Field of AI at EU Level*, in *Athena – Critical Inquiries in Law, Philosophy and Globalization*, 3(2), 2023, 135 ff., 148-149; G. Rugani, *La consultazione dei portatori di interessi nell'elaborazione degli atti dell'Unione Europea in materia di intelligenza artificiale: il caso dell'alleanza europea per l'IA*, in *Osservatorio sulle fonti*, 2, 2023, 407 ff., 422-423.

⁵⁹ European Commission, Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts, 21 April 2021, COM (2021) 206 final. On the AI Act Proposal, see for instance: M. Veale – F.Z. Borgesius, *Demystifying the Draft EU Artificial Intelligence Act*, in *Computer Law Review International*, 2021, 97 ff.; W.G. Voss, *AI Act: the European Union's Proposed Framework Regulation for Artificial Intelligence Governance*, in *Journal of Internet Law*, 2021, 7 ff.

⁶⁰ On the European Parliament's position, see for instance: J. Bracy, *European Parliament vote pushes AI Act significant step forward*, in *iapp.org*, 14 June 2023; K.J. Nahra and others, *European Parliament Adopts Negotiating Position on the AI Act*, in *WilmerHale Privacy and Cybersecurity Law Blog*, 15 June 2023; N. Aszódi, *EU Parliament vote on AI Act: Lawmakers chose to protect people against harms of AI systems*, in *algorithmwatch.org*, 15 June 2023; G. Rugani, *The European Parliament's Position on the AI Act: a human rights-oriented approach and an attempt to regulate generative AI systems*, in *Quaderni AISDUE – 2/2023*, Naples, 2023, 215 ff.

⁶¹ On the AI Act, see for instance: F. Ferri, *L'Unione europea e la nuova disciplina sull'intelligenza artificiale: questioni e prospettive* – *Rivista Quaderni AISDUE*, special issue no. 2/2024, Naples, 2024; C. N. Pehlivan – N. Forgó – P. Valcke, *The EU Artificial Intelligence (AI) Act: A Commentary*, Alphen aan den Rijn, 2025.

defence and national security purposes», the AI Act excludes from its application «AI systems where and in so far they are placed on the market, put into service, or used with or without modification exclusively for military, defence or national security purposes, regardless of the type of entity carrying out those activities» (art. 2, para. 3)⁶².

In regulating AI systems, the AI Act follows what is commonly referred to as a «risk-based approach»⁶³. Indeed, it differentiates between uses of AI that create an «unacceptable risk», a «high risk», and «low or minimal risk». AI systems that pose an unacceptable risk, exhaustively listed in art. 5 (Chapter II), are strictly prohibited: these are misuses of AI that provide for «manipulative, exploitative and social control practices», which are «particularly harmful and abusive» and «contradict Union values of respect for human dignity, freedom, equality, democracy and the rule of law and fundamental rights enshrined in the Charter, including the right to non-discrimination, to data protection and to privacy and the rights of the child»⁶⁴. The prohibition of unacceptable risk AI systems is already applicable since 2 February 2025, according to art. 113, lit. a).

The Regulation also identifies AI systems that pose a high risk of harm to the health and safety or the fundamental rights of natural persons, which are divided into two categories: AI systems intended to be used as safety components of products that are subject to third party *ex-ante* conformity assessment (art. 6, para. 1) and stand-alone AI systems with mainly fundamental rights implications (art. 6, para. 2), which are explicitly listed in Annex III of the Regulation. Such AI systems, in order to be placed on the Union market, put into service or used, must comply with stringent mandatory requirements⁶⁵, enshrined in Chapter III. In particular, Section 2 of such Chapter (articles 8-15) provides for risk management systems (art. 9), high quality of the datasets feeding the systems (art. 10), detailed documentation (art. 11), record-keeping (art. 12), transparency and clear information (art. 13), human oversight (art. 14), accuracy, robustness and cybersecurity (art. 15). Section 3 of Chapter III (articles 16-27), instead, is devoted to «Obligations of providers and deployers of high-risk AI systems and other parties»; the difference between provider and deployer is the following: the former is a natural or legal person, public authority, agency or other body that develops an AI system or that has it developed and places it on the market or puts it into service under its own name or trademark, whether for payment or free of charge (art. 3, n. 3); the latter is a natural or legal person, public authority, agency or other body using an AI system under its authority except where the AI system is used in the course of a personal non-professional activity (art. 3, n. 4). Among the responsibilities of the deployer, in certain cases, there is the obliga-

⁶² On the point see, for instance: D. Eder, *The AI Act and Bias. Effectiveness and Technical Feasibility of Countermeasures*, Linz, 2024; P. Vogiatzoglou, *The AI Act National Security Exception: room for manoeuvres?*, in *Verfassungsblog*, 9 December 2024.

⁶³ On the «risk-based approach», see for instance: G. De Gregorio - P. Dunn, *The European Risk-Based Approaches: Connecting Constitutional Dots in the Digital Age*, in *Common Market Law Review*, 59(2), 2022, 473 ff.

⁶⁴ Regulation (EU) 2024/1689, recital 28.

⁶⁵ *Ibid.*, recital 46.

tion to perform a «Fundamental rights impact assessment for high-risk AI systems» (Article 27), a topic to be further addressed below. With regard to Sections 4 (articles 28-39) and 5 (articles 40-49) of Chapter III, they concern «Notifying authorities and notified bodies» and «Standards, conformity assessment, certificates, registration», respectively.

Finally, AI systems whose risk level does not fall within the first two categories – thus considered as posing a low or minimal risk – are subject only, in certain cases, to transparency obligations (Chapter IV, i.e. art. 50). It is important to clarify that the provisions concerning high-risk as well as low or minimal risk AI systems will become applicable pursuant to art. 113 of the AI Act⁶⁶.

5. The strategies promoted by the AI Act to counter algorithmic discrimination, with particular reference to states of emergency

Having provided such an overview of the AI Act, it is now possible to focus more in particular on the provisions that are relevant for the purposes of the present analysis.

Preliminarily, a crucial element must be stressed: Regulation (EU) 2024/1689 contains no rules specifically addressing the use of AI during states of emergency⁶⁷. By contrast, there are several provisions that address the issue of algorithmic discrimination⁶⁸. Some of them, as will be shown, could be particularly relevant in states of emergency: of course, their effectiveness will have to be more precisely assessed when – once

⁶⁶ *Ibid.*, art. 113 («Entry into force and application»), according to which: «This Regulation shall enter into force on the twentieth day following that of its publication in the Official Journal of the European Union. It shall apply from 2 August 2026. However: (a) Chapters I and II shall apply from 2 February 2025; (b) Chapter III Section 4, Chapter V, Chapter VII and Chapter XII and Article 78 shall apply from 2 August 2025, with the exception of Article 101; (c) Article 6(1) and the corresponding obligations in this Regulation shall apply from 2 August 2027».

⁶⁷ Within the AI Act, the word «emergency» is used only in recital 58 and in Annex III, where it refers exclusively to «AI systems used to evaluate and classify emergency calls by natural persons or to dispatch or establish priority in the dispatching of emergency first response services, including by police, firefighters and medical aid, as well as of emergency healthcare patient triage systems» (which are classified as high-risk AI systems). Therefore, it does not appear to be a particularly relevant reference for our purposes. There are likewise no mentions of equivalent terms, such as «crisis».

⁶⁸ On how the issue of algorithmic discrimination is addressed in the AI Act, see for instance: L. Bosoer – M. Cantero Gamito – R. Rubio-Marin, *Non-discrimination and the AI Act*, in F. Pedrini (ed.), *Law in the age of digitalization*, Navarra, 2024, 85 ff; L. Deck – J.L. Müller – C. Braun – D. Zipperling – K. Niklas, *Implications of the AI Act for Non-Discrimination Law and Algorithmic Fairness*, Mainz, 2024; J. Ruohonen, *On Algorithmic Fairness and the EU Regulations*, in *arxiv.org*, 2024; K. Meding, *It's complicated. The relationship of algorithmic fairness and non-discrimination regulations in the EU AI Act*, in *arxiv.org*, 2025; A.J. Pagano – A. Cappiello – C. D'Elia, *AI Act Between Selectivity and Discrimination. The State of the Art of National and EU Regulation*, in *Business Law Review*, 46(4), 2025, 112 ff.; M.L. Rebrean – G. Malgieri, *Vulnerability in the EU AI Act: building an interpretation*, in *EAccT'25: Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency*, 2025, 1985 ff.

the AI Act has become fully applicable pursuant to its art. 113 – a genuine emergency situation arises, which would constitute a «stress test» for the resilience of the Regulation as a whole, much in the same way as the pandemic itself represented a «stress test» for Regulation (EU) 2016/679 (the General Data Protection Regulation or GDPR)⁶⁹. In the meantime, however, it is legitimate to seek to understand how the provisions of the AI Act might operate in practice.

Returning to the references contained in Regulation (EU) 2024/1689 regarding algorithmic discrimination, the latter is acknowledged as a matter of concern already in the recitals, where the following can be read: «AI systems [...] may lead to discrimination between persons or groups and may perpetuate historical patterns of discrimination, such as that based on racial or ethnic origins, gender, disabilities, age or sexual orientation, or may create new forms of discriminatory impacts»⁷⁰.

More importantly, when turning to the actual legally binding content of the AI Act, it should be highlighted that AI systems presenting the most severe risks of discrimination are either strictly prohibited or classified as high-risk and made subject to the corresponding obligations.

5.1. Prohibited AI systems

Let us begin with prohibited AI systems, which are such precisely because they contradict Union values, including «equality», and fundamental rights enshrined in the Charter, including the «right to non-discrimination» (as noted above)⁷¹. Such systems include, for example, those «for the evaluation or classification of natural persons or groups of persons over a certain period of time based on their social behaviour or known, inferred or predicted personal or personality characteristics», in certain cases (art. 5, para. 1, lit. c)⁷²: as was noted above with regard to mass surveillance and registration measures, commonly implemented during states of emergency, «AI systems providing social scoring of natural persons by public or private actors may lead to discriminatory outcomes and the exclusion of certain groups. They may violate the right to dignity and non-discrimina-

⁶⁹ On the point see, for instance: L. Bradford, M. Aboy, K. Liddell, *COVID-19 contact tracing apps: a stress test for privacy, the GDPR, and data protection regimes*, in *Journal of Law and the Biosciences*, 2020, 1 ff.; N. Vavoula, *The COVID-19 Pandemic as a Stress Test on the Right to Protection of Personal Data The Case of Greece*, in *Eu crim*, 16(2), 2021, 122 ff.

⁷⁰ Regulation (EU) 2024/1689, recital 58.

⁷¹ *Ibid.*, recital 28.

⁷² In particular, art. 5, para. 1, lit. c), of Regulation (EU) 2024/1689 prohibits «the placing on the market, the putting into service or the use of AI systems for the evaluation or classification of natural persons or groups of persons over a certain period of time based on their social behaviour or known, inferred or predicted personal or personality characteristics, with the social score leading to either or both of the following: (i) detrimental or unfavourable treatment of certain natural persons or groups of persons in social contexts that are unrelated to the contexts in which the data was originally generated or collected; (ii) detrimental or unfavourable treatment of certain natural persons or groups of persons that is unjustified or disproportionate to their social behaviour or its gravity».

tion and the values of equality and justice»⁷³.

The abovementioned category also includes AI systems «for making risk assessments of natural persons in order to assess or predict the risk of a natural person committing a criminal offence, based solely on the profiling of a natural person or on assessing their personality traits and characteristics», with some exceptions (art. 5, para. 1, lit. d)⁷⁴. As has been shown – in the case mentioned above concerning the algorithmic tool used in the United States during the CoViD-19 emergency to support decisions regarding which inmates would be released to home confinement – similar instruments may exhibit racial biases and produce serious discriminatory effects⁷⁵.

There are, however, other AI systems that are prohibited, among other reasons, because of their potential discriminatory effects. For example, those designed «to infer emotions of a natural person in the areas of workplace and education institutions», with some exceptions (art. 5, para. 1, lit. f)⁷⁶, since they «may lead to discriminatory outcomes and can be intrusive to the rights and freedoms of the concerned persons»⁷⁷; and also «the use of ‘real-time’ remote biometric identification systems in publicly accessible spaces for the purposes of law enforcement», with some exceptions (art. 5, para. 1, lit. h)⁷⁸, since they «can lead to biased results and entail discrim-

⁷³ Regulation (EU) 2024/1689, recital 31.

⁷⁴ In particular, art. 5, para. 1, lit. d), of Regulation (EU) 2024/1689 prohibits «the placing on the market, the putting into service for this specific purpose, or the use of an AI system for making risk assessments of natural persons in order to assess or predict the risk of a natural person committing a criminal offence, based solely on the profiling of a natural person or on assessing their personality traits and characteristics; this prohibition shall not apply to AI systems used to support the human assessment of the involvement of a person in a criminal activity, which is already based on objective and verifiable facts directly linked to a criminal activity».

⁷⁵ Recital 42 of Regulation (EU) 2024/1689 reads as follows: «In line with the presumption of innocence, natural persons in the Union should always be judged on their actual behaviour. Natural persons should never be judged on AI-predicted behaviour based solely on their profiling, personality traits or characteristics, such as nationality, place of birth, place of residence, number of children, level of debt or type of car, without a reasonable suspicion of that person being involved in a criminal activity based on objective verifiable facts and without human assessment thereof. Therefore, risk assessments carried out with regard to natural persons in order to assess the likelihood of their offending or to predict the occurrence of an actual or potential criminal offence based solely on profiling them or on assessing their personality traits and characteristics should be prohibited. In any case, that prohibition does not refer to or touch upon risk analytics that are not based on the profiling of individuals or on the personality traits and characteristics of individuals, such as AI systems using risk analytics to assess the likelihood of financial fraud by undertakings on the basis of suspicious transactions or risk analytic tools to predict the likelihood of the localisation of narcotics or illicit goods by customs authorities, for example on the basis of known trafficking routes».

⁷⁶ In particular, art. 5, para. 1, lit. f), of Regulation (EU) 2024/1689 prohibits «the placing on the market, the putting into service for this specific purpose, or the use of AI systems to infer emotions of a natural person in the areas of workplace and education institutions, except where the use of the AI system is intended to be put in place or into the market for medical or safety reasons».

⁷⁷ Regulation (EU) 2024/1689, recital 44.

⁷⁸ In particular, art. 5, para. 1, lit. h), of Regulation (EU) 2024/1689 prohibits «the use

inatory effects. Such possible biased results and discriminatory effects are particularly relevant with regard to age, ethnicity, race, sex or disabilities»⁷⁹. It may be summarised by noting that a significant number of AI systems prohibited under Regulation (EU) 2024/1689 are so classified, *inter alia*, precisely because they may produce discriminatory effects. Moreover, some of these prohibitions – such as those concerning some social scoring practices and certain decision-support tools – may prove particularly valuable even in preventing discrimination during states of emergency.

5.2. High-risk AI systems

Let us now take into consideration high-risk AI systems.

Firstly, it should be highlighted that Annex III of the AI Act designates as high-risk certain AI systems used in particularly sensitive areas, namely the following: 1) biometrics; 2) critical infrastructure; 3) education and vocational training; 4) employment, workers' management and access to self-employment; 5) access to and enjoyment of essential private services and essential public services and benefits; 6) law enforcement; 7) migration, asylum and border control management; 8) administration of justice and democratic processes. Some of these areas have already been mentioned among those in which the risks of discrimination are most severe, and this is precisely one of the reasons for their inclusion in the list.

Secondly, with regard to the obligations arising from the classification of AI systems as high-risk, it should be noted that some of them are specifically aimed at eliminating or at least mitigating the risks of discrimination. A relevant example is provided by art. 9, devoted to «Risk management systems». In particular, important in the context of non-discrimination is art. 9, para. 2, lit. a), which requires the identification and analysis of known and reasonably foreseeable risks that the high-risk AI system can pose to health, safety, or fundamental rights. As already explained, fundamental rights also include the right of non-discrimination. Thus, indirectly, art. 9 already requires the identification and mitigation of discrimination in high-risk AI systems⁸⁰.

The core of the non-discrimination regulation for high-risk systems, how-

of 'real-time' remote biometric identification systems in publicly accessible spaces for the purposes of law enforcement, unless and in so far as such use is strictly necessary for one of the following objectives: (i) the targeted search for specific victims of abduction, trafficking in human beings or sexual exploitation of human beings, as well as the search for missing persons; (ii) the prevention of a specific, substantial and imminent threat to the life or physical safety of natural persons or a genuine and present or genuine and foreseeable threat of a terrorist attack; (iii) the localisation or identification of a person suspected of having committed a criminal offence, for the purpose of conducting a criminal investigation or prosecution or executing a criminal penalty for offences referred to in Annex II and punishable in the Member State concerned by a custodial sentence or a detention order for a maximum period of at least four years».

⁷⁹ Regulation (EU) 2024/1689, 32.

⁸⁰ On the point see, for instance: F.Z. Borgesius – P. Hacker – N. Baranowska – A. Fabris. 2024. *Non-discrimination law in Europe, a primer. Introducing European non-discrimination law to non-lawyers*, in *arxiv.org*, 2024.

ever, is art. 10, devoted to «Data and data governance»⁸¹. More specifically, art. 10, para. 2, lit. f, requires that training, validation and testing datasets shall be subject to an «examination in view of possible biases that are likely to affect the health and safety of persons, have a negative impact on fundamental rights or lead to discrimination prohibited under Union law, especially where data outputs influence inputs for future operations». Such provision is completed by art. 10, para. 2, lit. g, which demands for «appropriate measures to detect, prevent and mitigate possible biases identified according to point (f)». Thus, art. 10 not only mandates the examination of biases that lead to discrimination, but also the mitigation of these biases⁸².

Moreover, there are also other relevant provisions, such as art. 15, devoted to «Accuracy, robustness and cybersecurity», para. 4 of which mandates that «High-risk AI systems that continue to learn after being placed on the market or put into service shall be developed in such a way as to eliminate or reduce as far as possible the risk of possibly biased outputs influencing input for future operations (feedback loops), and as to ensure that any such feedback loops are duly addressed with appropriate mitigation measures». In this case, the output of a system is regulated, in contrast to the input of an AI system in art. 10. The reasoning behind this is that AI systems should not become echo chambers for biases⁸³.

As previously noted, and as can be seen from the analysis of the provisions at stake, there are no specific obligations concerning algorithmic discrimination in states of emergency. However, the fact that the intensity of the abovementioned obligations depends – in line with the «risk-based approach» – on the level of bias risk presented by AI systems is certainly a useful element: it must indeed be assumed that in a situation of crisis, where such risks are exacerbated (as extensively discussed above), the obligations will become more stringent.

5.3. The fundamental rights impact assessment for high-risk AI systems

Particular attention should be accorded to the fundamental rights impact assessment (FRIA) for high-risk AI systems.

With regard to the origin of the FRIA, it was included neither in the Commission's Proposal nor in the Council's «General approach», but was introduced only by the European Parliament in art. 29a of its negotiating position. In particular, art. 29a, para. 1, of the EP's position stated that «Prior to putting a high-risk AI system as defined in Article 6(2) into use, with the exception of AI systems intended to be used in area 2 of Annex III, deployers shall conduct an assessment of the systems' impact in the specific context of use», including at minimum certain elements which are listed. Moreover, art. 29a, para. 4, of the EP's position provided as

⁸¹ K. Meding, *It's complicated*, cit., 9.

⁸² *Ibid.*, 10.

⁸³ *Ibid.*, 11.

follows: «In the course of the impact assessment, the deployer, with the exception of SMEs, shall notify national supervisory authority and relevant stakeholders and shall, to best extent possible, involve representatives of the persons or groups of persons that are likely to be affected by the high-risk AI system, as identified in paragraph 1, including but not limited to: equality bodies, consumer protection agencies, social partners and data protection agencies, with a view to receiving input into the impact assessment». Therefore, it was envisaged as mandatory to involve stakeholders in the FRIA.

In the final text of the AI Act, the FRIA is maintained, specifically regulated by art. 27, even though – following negotiations with the Council – its scope is now reduced compared with that in the EP's position. More specifically, art. 27, para. 1, states as follows: «Prior to deploying a high-risk AI system referred to in Article 6(2), with the exception of high-risk AI systems intended to be used in the area listed in point 2 of Annex III, deployers that are bodies governed by public law, or are private entities providing public services, and deployers of high-risk AI systems referred to in points 5 (b) and (c) of Annex III, shall perform an assessment of the impact on fundamental rights that the use of such system may produce», including certain elements which are listed. This means that the obligation to carry out the FRIA now applies only to deployers that are bodies governed by public law or private entities providing public services, and deployers of high-risk AI systems intended to be used to evaluate the creditworthiness of natural persons (Annex III, n. 5, lit. b) or for risk assessment and pricing in the case of life and health insurance (Annex III, n. 5, lit. c)⁸⁴.

Moreover, art. 27 no longer refers to the involvement of stakeholders in the FRIA. Such involvement is now only recommended in recital 96, according to which «to collect relevant information necessary to perform the impact assessment, deployers of high-risk AI system, in particular when AI systems are used in the public sector, could involve relevant stakeholders, including the representatives of groups of persons likely to be affected by the AI system, independent experts, and civil society organisations in conducting such impact assessments and designing measures to be taken in the case of materialisation of the risks»⁸⁵.

In any case, notwithstanding its limitations, the FRIA appears to be a mechanism that could enable a certain degree of dialogue with stakehold-

⁸⁴ Annex III of the AI Act reads as follows: «High-risk AI systems pursuant to Article 6(2) are the AI systems listed in any of the following areas: [...] Access to and enjoyment of essential private services and essential public services and benefits: [...] (b) AI systems intended to be used to evaluate the creditworthiness of natural persons or establish their credit score, with the exception of AI systems used for the purpose of detecting financial fraud; (c) AI systems intended to be used for risk assessment and pricing in relation to natural persons in the case of life and health insurance [...]».

⁸⁵ On the fundamental rights impact assessment (FRIA) for high-risk AI systems provided for in art. 27 of the AI Act, see for instance: A. Mantelero, *The Fundamental Rights Impact Assessment (FRIA) in the AI Act: Roots, legal obligations and key elements for a model template*, in *Computer Law & Security Review*, 54, 2024; M.E. Kaminski – G. Malgieri, *Impacted Stakeholder Participation in AI and Data Governance*, in *Yale Journal of Law & Technology*, 27(1), 2025, 247 ff.

ers. This is fully consistent with the fact that the AI Act aspires to represent not only a «point of arrival» of co-regulation in the field of AI – being the result of a long and articulated consultation (as mentioned above) – but also a «point of departure» for further processes of engagement with stakeholders, through several mechanisms⁸⁶. In particular, the FRIA could allow dialogue with the most vulnerable individuals and groups affected by AI systems and their discriminatory effects, in order to counter the latter.

It must never be forgotten that one of the reasons why algorithms discriminate is that those engaged in the field of AI are predominantly men, typically relatively young, white and Western (or, at most, Asian)⁸⁷. Other segments of society are instead underrepresented. First, among them are women, who are generally underrepresented in STEM (science, technology, engineering and mathematics) education and subsequent careers⁸⁸; a problem that is particularly acute in Western countries, as compared, for instance, with Arab and Islamic countries, where there are often higher rates of females pursuing STEM education⁸⁹. Greater involvement in the field of AI of underrepresented groups, through instruments such as the fundamental rights impact assessment, may therefore constitute a valuable strategy for better identifying potential biases in AI systems and tackling their discriminatory effects⁹⁰.

With particular reference to the FRIA, it may prove especially useful in countering algorithmic discrimination during states of emergency. As has been noted, following negotiations with the Council, the obligation to carry out the FRIA applies to a more limited number of deployers com-

⁸⁶ On the AI Act as a «point of arrival» and a «point of departure» of co-regulation in the field of AI, see in greater detail: G. Rugani, *La “legge sull’intelligenza artificiale” dell’UE come punto di arrivo e di partenza dei processi di co-regolazione*, in *Osservatorio sulle fonti*, 1, 2024, 509 ff.

⁸⁷ On the point see, for instance: J. Brown – T. Park – J. Chang – M. Andrus – A. Xiang – C. Custis, *Attrition of Workers with Minoritized Identities on AI Teams*, in *EAAMO ‘22: Proceedings of the 2nd ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization*, 2022.

⁸⁸ On the point see, for instance: C. Schmuck, *Women in STEM Disciplines*, Cham, 2017; S. Kahn – D. Ginther, *Women and Stem*, Cambridge, 2017; B.J. Casad – J.E. Franks – C.E. Garasky – M.M. Kittleman – A.C. Roesler – D.Y. Hall – Z.W. Petzel, *Gender inequality in academia: Problems and solutions for women faculty in STEM*, in *Journal of Neuroscience Research*, 99(1), 2021, 13 ff.

⁸⁹ On the point see, for instance: United Nations Educational, Scientific and Cultural Organization (UNESCO), *Women in Science. Fact Sheet No. 51*, June 2018; S.I. Islam, *Science, Technology, Engineering and Mathematics (STEM): Liberating Women in the Middle East*, in *World Journal of Education*, 9(3), 2019, 94 ff.; R. Dajani – S. Dhawan – S. M. Awad, *The Increasing Prevalence of Girls in stem Education in the Arab World. What Can We Learn?*, in *Sociology of Islam*, 8, 2020, 159 ff.

⁹⁰ On the usefulness of stakeholder involvement in countering algorithmic discrimination, see for instance: G. Cachat-Rosset – A. Klarsfeldb, *Diversity, Equity, and Inclusion in Artificial Intelligence: An Evaluation of Guidelines*, in *Applied Artificial Intelligence*, 37(1), 2023; J. Solyst – E. Yang – S. Xie – A. Ogan – J. Hammer – M. Eslami, *The Potential of Diverse Youth as Stakeholders in Identifying and Mitigating Algorithmic Bias for a Future of Fairer AI*, in *Proceedings of the ACM on Human-Computer Interaction* 7 CSCW2, 2023; M.E. Kaminski – G. Malgieri, *Impacted Stakeholder Participation*, cit.

pared to the European Parliament's original position. Nevertheless, the categories of deployers for which the FRIA has remained mandatory in the final version of the AI Act are precisely those playing the most significant role in situations of crisis: we refer in particular to bodies governed by public law (and private entities providing public services), which have primary responsibility in cases such as those discussed above concerning mass surveillance and registration measures and decision-support tools. For the purpose of countering algorithmic discrimination in emergencies, the fact that such bodies and entities are required to carry out the FRIA must therefore be welcomed.

6. Concluding remarks

It is now useful to summarise what this contribution has shown so far. As has long been widely demonstrated, AI can produce discriminatory effects against women, as well as various minority and vulnerable groups (e.g., individuals belonging to racial or ethnic minorities, persons with disabilities, people who are not heterosexual or cisgender). This depends on several factors, the most significant of which is «data training», since the more biased or prejudiced the data are, the more the model will reflect such bias, acting in a discriminatory manner. In any case, algorithms which perpetuate or amplify existing forms of discrimination inevitably contribute to reinforcing social inequalities across a wide range of contexts.

During states of emergency, moreover, the state-business complex tends to increase its reliance on digitalised and AI-based solutions, a circumstance that significantly intensifies both the likelihood of algorithmic discrimination and the associated risks: on the one hand, the increase in the deployment of technologies capable of producing discriminatory outcomes heightens the probability of such effects; on the other hand, during crises these technologies are especially applied in highly sensitive contexts, where the consequences of discriminatory decisions can be particularly severe.

The paper, in particular, took into consideration two of the main categories of digitalised and AI-based solutions most frequently employed during states of emergency: on the one hand, mass surveillance and registration measures; on the other, decision-support tools. It clearly emerged that the widespread use of similar solutions, which is typically carried out in response to crises, has the undeniable potential to deepen discrimination against already marginalised groups.

Given the above, this contribution has aimed to answer a specific research question: whether, and to what extent, the AI Act – the content of which has been duly summarised – attempts to promote strategies to counter algorithmic discrimination, with particular reference to that occurring in states of emergency.

Following the analysis undertaken, it can be stated that – although Regulation (EU) 2024/1689 contains no rules specifically addressing the use of AI during states of emergency – there are several provisions that address the issue of algorithmic discrimination, some of which could be particu-

larly relevant in states of emergency.

With reference to prohibited AI systems (art. 5), in many cases they are classified in such category, *inter alia*, precisely because they may produce discriminatory effects. Some of these prohibitions, moreover, may prove particularly valuable even in preventing discrimination during states of emergency: for example, those concerning some social scoring practices and certain decision-support tools, which (as has been highlighted) have often been relied upon more heavily during crises.

As for high-risk AI systems, they require compliance with various obligations, some of which are specifically aimed at eliminating or at least mitigating the risks of discrimination (e.g., articles 9, 10 and 15). Furthermore, since the intensity of such obligations depends on the level of bias risk presented by AI systems, it must be assumed that in a situation of crisis (where similar risks are exacerbated) the obligations will become more stringent.

Still concerning high-risk systems, a particularly useful tool is the fundamental rights impact assessment (FRIA, art. 27). The latter applies to a limited number of deployers; however, these include those bearing the greatest responsibilities during emergencies, notably bodies governed by public law (and private entities providing public services). With respect to these deployers, the FRIA could allow dialogue with the most vulnerable individuals and groups affected by AI systems and their discriminatory effects, and help to counter these effects.

However, although the elements of the AI Act discussed so far should be regarded as steps forward compared to the pre-existing legal framework, they do not yet appear entirely satisfactory.

Firstly, the provisions of the AI Act aimed at countering algorithmic discrimination, although present, lack specificity: only arts. 10 and 15 explicitly mention non-discrimination and biases, but still without dealing extensively with the issue, while in art. 9 there is just a vague statement concerning fundamental rights protection⁹¹. Equally concerning is the complete lack of *ad hoc* rules regarding states of emergency (as repeatedly highlighted): in light of the significant risks that the use of AI systems in these contexts can pose, a less fragmented and more precise legal framework on the matter would be desirable.

Secondly, regarding the main instrument considered in this contribution, namely the FRIA, there are strong reasons to believe that the amendments to the European Parliament's original position, introduced during the negotiations with the Council, may have significantly weakened the instrument. In particular, the fact that stakeholder involvement is no longer mandatory – as it was under art. 29a, para. 4, of the EP's position – but is merely recommended in recital 96, could seriously undermine the usefulness of the fundamental rights impact assessment as a tool to dialogue with vulnerable individuals and groups in order to counter algorithmic discrimination. Moreover, in any case, instruments such as the FRIA involve stakeholder engagement only at the final stage, just before systems are placed on the market; instead, it would arguably be worth considering

⁹¹ K. Meding, *It's complicated*, cit., 12.

mechanisms that allow their participation at an earlier stage, namely during the design and development of AI systems.

Thirdly, as a side note, it should not be entirely overlooked that some of the exceptions to the scope of application of the Regulation also raise concerns, as they could loosen constraints on fundamental rights: we refer, for instance, to the one concerning AI systems for military, defence and national security purposes, given that such fields could well come into play during crises.

In conclusion, the AI Act undoubtedly represents a first step at EU level towards recognising the challenges posed by algorithmic discrimination, and it provides some initial tools to counter it, certain of which may prove particularly useful in states of emergency. In the near future, however, further steps should be taken, with the aim of establishing a more specific and detailed framework and of fostering, as far as possible, the precious dialogue with stakeholders. Moreover, the effectiveness of the existing provisions will need to be assessed in the context of a genuine emergency, which would constitute a true «stress test» of the AI Act's overall resilience.

Abstract

This contribution focuses on Regulation (EU) 2024/1689, better known as «AI Act», in order to address a specific research question: whether, and to what extent, the AI Act attempts to promote strategies to counter algorithmic discrimination, with particular reference to that occurring in states of emergency. To this end, the paper preliminarily outlines the issue of algorithmic discrimination in general, identifying in brief the most affected categories, the reasons why AI systems can produce discriminatory effects, and the resulting consequences. Then, it analyses the specific matter of algorithmic discrimination in states of emergency, explaining why such phenomenon raises particular concerns: for this purpose, it considers two of the main categories of digitalised and AI-based solutions most frequently employed during crises – i.e., mass surveillance and registration measures, and decision-support tools – assessing their potential to produce discriminatory effects. At that point, the contribution turns to the AI Act, examining the presence of provisions aimed at countering not only algorithmic discrimination in general, but also that which can occur in states of emergency. Finally, some concluding remarks are offered, highlighting not only the positive aspects, but also the limitations and margins for improvement of the current EU legal framework.

Il presente contributo si concentra sul Regolamento (UE) 2024/1689, meglio noto come «AI Act», al fine di rispondere a una specifica domanda di ricerca: se, e in quale misura, l'AI Act tenta di promuovere strategie per contrastare la discriminazione algoritmica, con particolare riferimento a quella che si manifesta negli stati di emergenza. A tal fine, l'articolo delinea preliminarmente il tema della discriminazione algoritmica in generale, individuando in breve le categorie più colpite, le ragioni per cui i sistemi di IA possono produrre effetti discriminatori e le conseguenze che ne derivano. Poi, viene analizzato il problema specifico della discriminazione algoritmica negli stati di emergenza, illustrando le ragioni per cui tale fenomeno solleva particolari preoccupazioni: a tal fine, vengono considerate due delle principali categorie di soluzioni digitalizzate e basate sull'IA più frequentemente impiegate durante le crisi – cioè le misure di sorveglianza e registrazione di massa e gli strumenti di supporto decisionale – valutandone l'idoneità a produrre effetti discriminatori. A questo punto, il contributo si dedica all'AI Act, esaminando la presenza di disposizioni volte a contrastare non solo la discriminazione algoritmica in generale, ma anche quella che può verificarsi negli stati di emergenza. Infine, vengono formulate alcune considerazioni conclusive, che evidenziano non solo gli aspetti positivi, ma anche i limiti e i margini di miglioramento dell'attuale disciplina giuridica UE.

Keywords

regulation (EU) 2024/1689 (AI Act) - artificial intelligence - algorithmic discrimination - state of emergency - stakeholder consultation

Saggi

The Brussels Sphinx's Riddle. What is a high-risk AI System?*

Andrea Bertolini, Federica Fedorczyk, Marta Mariolina Mollicone, Guilherme Migliora**

Table of Contents

1. Introduction. – 2. The Morning - The regulatory framework – 2.1. The AIA as a New Piece of Product Safety Legislation. – 2.2. The risk-based approach as a way to offset the horizontal approach of the AIA. – 2.3. The horizontal importance of definitions. – 3. The Midday - Art. 6 and the definition of h-AIS: Annex III and the high-risk categories. – 3.1. The reference to annex III. – 3.2. Exceptions and future amendments. – 3.3. Amendments to annex III. – 3.4. The “self-declaration”. – 4. The Evening - Art. 6, para. 1, of the AIA: defining h-AIS. – 4.1. The complexity of para. 1. – 4.2. The safety component. – 4.3. Safety components and complex systems. – 4.4. The certification of the AIS or of its safety component: an overview. – 4.5. When is a TPCA “required”? – 4.6. Case studies: on the marginal discrepancies – 5. A critical perspective. – 6. The unsolved riddle of art. 6 AIA.

1. Introduction

The Sphinx of Thebes, a creature with the body of a lion, the wings of an eagle, and the face of a woman, is a figure from Greek mythology that has been long discussed by scholars. It is notable for its tendency to pose riddles to passersby, which were rich in symbolism. Those who were unable to answer correctly were killed and devoured by the Sphinx.

In the riddle, the Sphinx inquired as follows:

«What is the creature that has only one voice, but walks on four legs in the morning, on two at noon, and on three in the evening?»¹

The answer was a human. Indeed, the morning represents the childhood

* L'articolo è stato sottoposto, in conformità al regolamento della Rivista, a referaggio “a doppio cieco”.

** Andrea Bertolini, Ph.D., LL.M. (Yale) is Associate Professor of Private Law at Scuola Superiore Sant'Anna (SSSA) and Director of The European Center of Excellence on The Regulation of Robotics And AI (EURA), supervised the structure of the entire article, and is primarily responsible for §§2.2-2.3, 4.2-4.3, 4.5, and 6; Federica Fedorczyk, Ph.D. is post-doctoral research fellow at Oxford University, is primarily responsible for §§2.2-3.4; Marta Mariolina Mollicone, Ph.D. is post-doctoral researcher at SSSA, and is primarily responsible for §§3.2 and 4.5-4.6; Guilherme Migliora, is Ph.D. candidate at SSSA and is primarily responsible for §§4.1-4.5. All websites were accessed no later than August 1st 2025.

¹ Apollodoro (Pseudo), *Biblioteca*, 3, 5, vv. 7-8.

where a baby crawls on all fours, using four limbs, two hands and two knees; the noon is the adulthood. An adult walks on two legs, upright. The evening refers to old age since an elderly person uses a cane, so they “walk on three legs”, two legs plus a cane.

Ultimately that of the Sphinx was a question of classification, not too different from that the European legislator – who assumes the role of the giant regulator of artificial intelligence (henceforth AI) – poses to those who approach the development of a complex AI system (henceforth AIS). Indeed, the riddle they face is the accurate classification of their system pursuant to art. 6 of the AI Act (henceforth AIA)², and if the Brussels’ Sphinx where to phrase it, it could be as follows:

“What is the AIS that falls under Annex III, and yet does not fall under the exemption of paragraph 3, or at least performs some profiling function, and is not declared by its provider not to be high-risk pursuant to paragraph 4 or, instead, is itself, or one of its safety components, subject to the product safety legislation enumerated by Annex I, and itself or its safety component, require a third party conformity assessment to be certified for a lawful distribution on the European market?”

Most likely the Sphinx of Thebes would herself be confused when formulating such an articulate question; maybe that of Brussels, instead, is better trained in both tongue- and mind-twisters and would thence find herself at ease laying the question down.

Unlike Oedipus, however, providers and deployers of AI might not find themselves triumphant, and while the repercussions of error are not as dire as in the case of death, they are nevertheless grave.

Indeed, on the one hand, both the classification of a system as high risk and its erroneous designation as such entail the possibility of being subject to the penalties provided for in art. 99 et seq. AIA – in the physiology of technological development and adoption – if a system is qualified as high-risk providers and deployers³ will need to comply with an articulate series of requirements that apply across all phases of their lifecycle, from design and development to production, deployment, and even the so-called post-contractual phase, after the systems have been distributed and are operating on the market (art. 8-27 AIA). In such a perspective, the more conservative approach of considering a system as high-risk when-

² European Parliament and the Council, Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act), OJ L, 2024/1689, 12.7.2024.

³ Pursuant to art. 3, para. 3, AIA a provider «means a natural or legal person, public authority, agency or other body that develops an AI system or a general-purpose AI model or that has an AI system or a general-purpose AI model developed and places it on the market or puts the AI system into service under its own name or trademark, whether for payment or free of charge». Pursuant to art. 3, para. 4, deployer «means a natural or legal person, public authority, agency or other body using an AI system under its authority except where the AI system is used in the course of a personal non-professional activity».

ever in doubt, would represent a very burdensome strategy, and one that could come at a relevant economic and strategic cost, in particular with respect to possible competition.

Similarly to the fated heir of the Labdacids, we will attempt to solve the riddle, by deconstructing the articulate definition of High-Risk AI System (henceforth h-AIS) as put forth by art. 6 AIA. The awaited prize is the dominion over the citadel of European AI regulation. Indeed, the AIA horizontal regulatory approach⁴ (see §3.1), intended to be “future proof”⁵, is partially tempered by a risk-based categorization⁶, that causes the very

⁴ The AIA is a horizontal legislation as it creates a supposedly far-reaching framework for AI regulation across various sectors and applications, while attempting to balance the numerous risks and benefits AI can provide. It has the specific objective, among others, to ensure that AIS placed, put into service and used in the Union market are safe and respectful of existing law and fundamental rights. To this end, the AIA explicitly recalls the Ethics Guidelines for Trustworthy AI (EGTAI) developed by the independent High-Level Expert Group on AI (AI HLEG) appointed by the European Commission in 2018. The EGTAI had identified seven non-binding ethical principles, to be taken into account to achieve the design of a “human-centric” AI (i) Human agency and oversight: AIS should be developed and used as a tool that serves people, respects human dignity and personal autonomy, and that is functioning in a way that can be appropriately controlled and overseen by humans; ii) Technical robustness and safety: AIS should be developed and used in a way that allows robustness in the case of problems and resilience against attempts to alter the use or performance of the AI system so as to allow unlawful use by third parties and minimize unintended harm; iv) Privacy and data governance: AIS should be developed and used in accordance with privacy and data protection rules, with processing data that meets high standards in terms of quality and integrity; iv) Transparency: AIS should be developed and used in a way that allows appropriate traceability and explainability, while making humans aware that they communicate or interact with an AIS, as well as duly informing deployers of the capabilities and limitations of that AIS and affected persons about their rights; v) Diversity, non-discrimination and fairness: AIS should be developed and used in a way that includes diverse actors and promotes equal access, gender equality and cultural diversity, while avoiding discriminatory impacts and unfair biases that are prohibited by Union or national law; vi) Social and environmental well-being: AIS should be developed and used in a sustainable and environmentally friendly manner as well as in a way to benefit all human beings, while monitoring and assessing the long-term impacts on the individual, society and democracy). Those same principles are today recalled, without prejudice to the legally binding requirement of the AIA itself (and of any relevant applicable Union law), as criteria that are to be considered in the design and use of AI models. However, the requirements laid down by the AIA are substantially more stringent and immediately applicable across all Member States. Indeed, as a regulation, the AIA has direct and immediate applicability across all EU member states, meaning it becomes binding in its entirety without the need for any additional national legislation or acts of implementation. This uniform application ensures consistency in the legal framework governing AI throughout the Union, reducing the risk of fragmentation and promoting harmonized standards. In such a perspective, it certainly reflects a commitment to uphold the digital sovereignty of the European Union. For the announcement of the AI Act refer to: European Commission, *European Artificial Intelligence Act comes into force*, 2021, in ec.europa.eu.

⁵ European Commission, *Artificial Intelligence – Questions and Answers*, in ec.europa.eu.

⁶ The pyramid of risks distinguishes unacceptable (art. 5 AIA), high (art. 6 AIA and following), limited (art. 50 AIA), and minimal or no risk, devising radically different regulatory frameworks. A separate section is dedicated to general-purpose AI (GPAI) models, which must comply with specific obligations related to intellectual property protection, transparency, and the mitigation of systemic risks (art. 53–55 AIA) that, however, entirely fall outside the scope of this study.

definition of h-AIS to become of paramount importance, for it determines the scope of application of those provisions which also represent the most relevant core of the entire regulation.

From a practical perspective, this will provide providers with the necessary guidance to understand how their AIS may qualify, and subsequently what obligations they must comply with. From a theoretical perspective, it allows for a discussion of the effectiveness of the emergent framework, its ambiguities, and the interpretive challenges that fraught it. Ultimately, it will increase awareness about the complexity of the application of the AIA, and about the need to collaborate with lawyers, already in the design phase, to ensure compliance with EU regulation.

Much like the original riddle – where the day is divided into three parts to represent the stages of human life: childhood, adulthood, and old age – this analysis is structured in three corresponding sections. “The Morning” sets the stage by outlining the background of the AIA and the principles on which is grounded, with particular attention to its risk-based and horizontal regulatory approaches. It also explores the Act’s broad definition of AIS and its commitment to technological neutrality, using concrete examples to illustrate the artificial nature of risk categorization across varied applications.

“The Midday” turns to art. 6, para. 2, examining the criteria for what qualifies as “high-risk” AIS. This section unpacks the structure of art. 6 and its reference to Annex III, revealing inconsistencies in the classification of high-risk applications.

“The Evening” focuses into art. 6, para. 1, analysing the criteria used to identify high-risk systems and highlighting the conceptual and linguistic shortcomings of this provision, along with its implications for both providers and deployers. The paper concludes summarizing the key findings in tables for comparative insights.

2. The Morning – The Regulatory Framework

2.1. The AIA as a New Piece of Product Safety Legislation

On July 12, 2024, the European Union marked a significant milestone by approving the AIA, the world’s first comprehensive regulation of artificial intelligence. It is important to note that while the AIA is formally a regulation, it represents a rather distinctive and unconventional type. Although it aims for maximum harmonization across the EU, it simultaneously leaves significant room for interpretation and flexibility. The Act calls for numerous delegated acts⁷ and grants substantial discretion to member states,

⁷ The power to adopt delegated acts is referred to in art. 6, para. 6 and 7, to add new conditions to those laid down in the art., by amending para. 3; art. 7, para. 1 and 3, allows for the Commission to add, modify or remove use-cases of h-AIS; art. 11, para. 3, refers to the possibility of amending Annex IV in regard to the necessary technical documentation; art. 43, para. 5 and 6 allows for the amendment of conformity assessment requirements in para. 1 and 2 of the same art. and Annexes VI and VII; art. 47, para.

allowing them considerable leeway in interpreting and adapting its provisions. For instance, Member States (henceforth MS) can, in exceptional cases (e.g. public security or protection of life and health of natural persons, environmental protection and the protection of key industrial and infrastructural assets), authorize the placing on the market or the putting into service of AI systems, which have not undergone a conformity assessment art. 46 and Recital 130 AIA). This not only creates a high degree of complexity but also raises fundamental questions about the very nature of this legislative instrument. One could even question whether the AIA truly functions as a regulation in the strict sense, or whether it shares characteristics more akin to a directive, ultimately blurring the traditional line between these two forms of EU legislation. Moreover, when we turn to its substance, the AIA is, at its core, primarily a piece of product safety legislation⁸.

Despite being a complex and highly articulated discipline, the AIA can, with some approximation, be qualified as a new example of product safety legislation⁹, primarily on the basis of its core obligations (see art. 5, 6 and 40-49) that establish a regulatory structure heavily reliant on risk regulation and certification through conformity assessment¹⁰.

To better understand its structure, it is therefore necessary to clarify some fundamental characteristics of European product safety legislation, rooted in the so-called “New Approach” and the subsequent New Legislative Framework (henceforth NLF)¹¹, expressly referenced to by Recital 9 of the AIA.

European product safety legislation, comprising over 20 directives and regulations¹², only establishes essential safety requirements, specific to different classes of applications, products, devices and components, leaving it to manufacturers to determine how to meet them. An example of such freedom granted to manufacturers is provided by art. 10, para. 1, and Annex III of the forthcoming Machinery Regulation (MR)¹³. Manufacturers

5, allows the Commission to update the content of the EU declaration of conformity set out in that Annex V; art. 51, para. 3, through which the Commission can amend the threshold for classifying a GPAI as a model with systemic risk; art. 52, para. 4, allows the Commission to update the criteria that justify a qualified alert from the scientific panel (Annex XIII); and art. 53, para. 5 and 6, allows for the Commission to detail measurement and calculation methodologies referred in the art. and to amend Annexes XI and XII.

⁸ *Ex multis*, A. Iannuzzi, *Le fonti del diritto dell'Unione europea per la disciplina della società digitale*, in F. Pizzetti – M. Orofino – E. Longo – A. Iannuzzi – S. Calzolaio (eds.), *La regolazione europea della società digitale*, Torino, 2024, 9 ff.

⁹ M. Almada – N. Petit, *The EU AI Act: a medley of product safety and fundamental rights?*, in *Robert Schuman Centre for Advanced Studies Research Paper No. 2020/14*, 2023, 7 ff.

¹⁰ P. Hacker, *AI Regulation in Europe: From the AI Act to Future Regulatory Challenges*, in *arXiv.org*, 2023, 1 ff.

¹¹ European Commission, *New legislative framework*, in *single-market-economy.ec.europa.eu*.

¹² The AIA, in its Annex 1, contains a list of Union harmonisation legislation based on the New Legislative Framework and a list of List of other Union harmonisation legislation.

¹³ European Parliament and the Council, Regulation (EU) 2023/1230 of the European Parliament and of the Council of 14 June 2023 on Machinery and Repealing Directive

must comply with the rules provided by the MR, such as that a machinery must «take the necessary protective measures in relation to risks that cannot be eliminated». However, the Regulation leaves room for the manufacturers to decide the means through which this requirement will be achieved. This gives the manufacturers more freedom to innovate.

The regulatory frameworks are then complemented by a very broad, articulate and distinct set of standards, often referred to by engineers as “norms”¹⁴, thus potentially leading to ambiguity about their legal value. Indeed, while standards certainly play an essential role in promoting safety and interoperability, they are never binding. That entails that when developing any application requiring to be certified pursuant to one or more pieces of European product-safety legislation, the party is only bound to comply with the corresponding directive(s) and regulation(s), never with a technical standard, irrespective of the authority having adopted them. Most technical standards are, in fact, developed at the international level by organizations such as ISO (International Organization for Standardization)¹⁵, IEC (International Electrotechnical Commission)¹⁶, and IEEE (Institute of Electrical and Electronics Engineers)¹⁷. At the European level, the so-called European Standardization Organizations (henceforth ESOs) – the European Committee for Standardization (CEN), the European Committee for Electrotechnical Standardization (CENELEC) and the European Telecommunications Standards Institute (ETSI) – operate, developing standards applicable within MS. However, relevant efforts are made to ensure alignment with the international framework to avoid market fragmentation, in particular through a series of agreements¹⁸ that ensure coordination between the standardization efforts initiated with a ESO and their corresponding international equivalent¹⁹.

2006/42/EC of the European Parliament and of the Council and Council Directive 73/361/EEC, OJ L 165, 29.6.2023.

¹⁴ For instance, in Germany and Italy the terminology used in the standardization systems - *normen* in German, issued by the *Deutsches Institut für Normung* (DIN) and *norme* in Italian, issued by the Ente Nazionale Italiano di Normazione (UNI) – does not imply that these are legally binding per se. However, such norms often hold considerable authority in engineering and technical fields, and in some cases, they are formally integrated into legal regulations.

¹⁵ ISO, *ISO - About ISO*, in *iso.org*.

¹⁶ IEC, *IEC Webstore Homepage*, in *iec.ch*.

¹⁷ IEEE Standards Association, *IEEE Standards Association*, in *standards.ieee*

¹⁸ CEN has an agreement for technical co-operation with the International Organization for Standardization (ISO), called the Vienna Agreement, signed in 1991. CENELEC enjoys close cooperation with its international counterpart, the International Electrotechnical Commission (IEC), also thanks to the Frankfurt Agreement of 2016. Unlike CEN and CENELEC, ETSI does not rely on a single equivalent “agreement” akin to the Vienna or Frankfurt Agreements. Instead, ETSI’s cooperation is structured through a variety of memorandum of understanding (MoUs) and partnership arrangements, which can be accessed at the following link portal.etsi.org.

¹⁹ Under the Vienna Agreement, if CEN begins developing a standard and it is determined that ISO is already working on the same subject, CEN suspends its development activities. This coordination avoids duplication of efforts and promotes the adoption of international standards at the European level. For example, CEN will stop its drafting work and monitor ISO’s progress, potentially adopting the final ISO standard

The most distinctive trait of European standardization is instead represented by the so-called harmonized standards (henceforth hENs)²⁰, whose importance is expected to grow with the AIA, especially since the European Commission (henceforth EC) has begun to promote the adoption of many such standards to address evident gaps related to more advanced applications, falling within the scope of the AIA²¹. European standardization organizations, led by CEN and CENELEC, are in the process of drafting the necessary AI standards, following a request from the EC²². The main advantage ensured by this very special category of standards is that adherence to them ensures a rebuttable presumption of conformity with the corresponding piece of product safety legislation.

Overall, a regulatory technique, heavily reliant upon standards, may elicit more than a subtle criticism in light of the lack of democratic control and accountability – primarily – typical of transnational private regulators²³,

as a European Standard (EN ISO). This procedure is detailed in CEN, *Guidelines for the implementation of the Vienna Agreement*, in boss.cen.eu.

²⁰ See European Parliament and the Council, Regulation (EU) No 1025/2012 of the European Parliament and of the Council on European standardization, OJ L 316, 14.11.2012. Art. 2, para.1, lit. c) defines a «harmonized standard» as «a European standard adopted on the basis of a request made by the Commission for the application of Union harmonization legislation». Among the other relevant pieces of legislation, see European Parliament and the Council, Regulation (EC) No 765/2008 of the European Parliament and of the Council of 9 July 2008 setting out the requirements for accreditation and market surveillance relating to the marketing of products, OJ L 218, 13.8.2008, and European Parliament and the Council, Decision No 768/2008/EC of the European Parliament and of the Council of 9 July 2008 on a common framework for the marketing of products, OJ L 218, 13.8.2008. Consider also the EC documents, especially European Commission, Communication of 1 June 2011, *A strategic vision for European standards: Moving forward to enhance and accelerate the sustainable growth of the European economy by 2020*, COM(2011) 311 final; European Commission, Communication of 1 June 2016, *European standards for the 21st century*, COM(2016) 358 final; and European Commission, Communication of 22 November 2018, *The Single Market in a changing world: A unique asset in need of renewed political commitment*, COM(2018) 772 final. In this last document, the EC states that «standardization has played a leading role in developing the Single Market by supporting market-based competition and by helping ensure the interoperability of products and services.

Products which comply with voluntary, harmonized standards endorsed at Union level benefit from a presumption of conformity and can therefore move freely in the Single Market. This has been very beneficial, for instance, in the field of engineering or information technologies. While harmonized standards are developed by European standardization bodies, the Commission initiates, manages and monitors these standards and bears the ultimate responsibility» as stated by the Court of Justice of the European Union (ECJ). See also European commission, Communication of 22 November 2018, *Harmonized standards: Enhancing transparency and legal certainty for a fully functioning Single Market*, COM(2018) 764 final, where the Commission recognized certain shortcomings of the EU standardization process —particularly in the context of rapidly evolving technological developments—and declared its commitment to addressing them.

²¹ European Commission, *Commission Implementing Decision C(2023) 3215 final on a standardisation request to the European Committee for Standardisation and the European Committee for Electrotechnical Standardisation in support of Union policy on Artificial Intelligence*, Brussels, 2023, in europa.eu

²² The first request for standards was already advanced by the EC to ESOs, see European commission, *Harmonised Standards for the European AI Act*, in publications.jrc.ec.europa.eu.

²³ Indeed, many of these standards are developed with a strong influence of private

especially when such technical elements may impinge upon the fundamental rights of the individual. Indeed, one of the main aspects differentiating the AIA from other pieces of—pure—product safety legislation is represented by its focus on the potential violation of fundamental rights²⁴, and on the possible tension that may arise when technical standards interact with the protection of those rights²⁵.

2.2. The Risk-based Approach as a Way to Offset the Horizontal Approach of the AIA

While the notion of risk is clearly central in the AIA²⁶, its concept may

interests often outweighing public accountability, raising worries about the limited democratic oversight in the standard-setting process. See C. Scott – F. Cafaggi – S. Linda, *The Conceptual and Constitutional Challenge of Transnational Private Regulation*, in *Journal of Law and Society*, 38, 2011, 1 ff. Furthermore, in such a perspective, it shall be noted that this reliance on privately developed standards triggers serious concerns about transparency, accountability, and potential conflicts of interest. Because private stakeholders dominate much of the standardization process, crucial product safety rules are shaped largely outside democratic institutions. For this reason, the ECJ addressed this issue in its landmark 2013 ruling (Case C-613/12, *James Elliott Construction Limited v. Irish Asphalt Limited*), asserting that harmonized standards fall within the scope of EU law and are therefore subject to judicial review. This ruling underscored the principle that, although technical standards are developed by private entities, they remain subject to some form of democratic scrutiny. The decision has since been widely discussed in legal scholarship, with many commentators arguing that it represents a necessary check on the growing influence of private standard-setting bodies in public regulatory matters. See, for instance: R. Vallejo, *The Private Administrative Law of Technical Standardization*, in *Yearbook of European Law*, 40, 2021, 172 ff; C. Tovo, *Judicial review of harmonized standards: Changing the paradigms of legality and legitimacy of private rulemaking under EU law*, in *Common Market Law Review*, 55, 2018, 1187 ff.

²⁴ This focus is also specifically recalled by Recitals 6 and 7 AIA, especially when referring to h-AIS. M. Almada – N. Petit, *The EU AI Act: a medley of product safety and fundamental rights?*, cit., 7 ff.

²⁵ In this sense, see L. Lessing, *Code Version 2.0*, 2006, in which the author conceptualizes technical architectures as regulatory structures that shape and constraint behaviour in ways comparable to legal rules. When applied to standardization processes, this underscores how technical norms drafted by private or transnational bodies may predetermine how legal rights are exercised in practice. Such *de facto* regulatory effects may generate tensions with fundamental rights guarantees, insofar the design choices embedded in technical norms can restrict or privilege certain rights.

²⁶ Observing current global regulatory initiatives for AI, it is evident that the risk-based approach appears to be dominant in a global perspective when considering AI regulation. For example, the Bletchley Declaration, signed in November 2023, by 28 countries (including the United States, China, and the European Union) during the UK AI Safety Summit, emphasizes that nations should account for AI-related risks and, where appropriate, adopt «classifications and categorizations of risk based on national circumstances and applicable legal frameworks». Similarly, the Council of Europe's Convention on Artificial Intelligence integrates general principles with a risk-based approach. It mandates actions for «identification, assessment, prevention, and mitigation of risks and impacts to human rights, democracy, and the rule of law» at all stages of the AI lifecycle, from design to decommissioning. Meanwhile, in the United States, President Biden's Executive Order on AI assesses risks associated with dual-use foundation models, especially those with publicly available model weights, and explores potential mechanisms

appear ambiguous. Indeed, risk-based regulation is typically non-European, and relies upon the exact assessment of an existing risk, possibly measured and quantified by scientific research²⁷. To the contrary, the regulation of precautionary principle has always characterized European legislation²⁸, requiring the adoption of measures even when scientific evidence remains inconclusive or uncertainty persists²⁹, so long as there are reasonable grounds for concerns³⁰.

In such a perspective, the AIA's risk-based approach emphasizes proactive risk mitigation aiming to prevent potential harm before it materializes, rather than merely reacting to identified risks, and may thence be rightfully framed within the European tradition, rooted in the precautionary principle³¹.

to mitigate risks while maximizing societal benefits.

²⁷ In contrast, the risk-based approach employed in the United States often operates *ex post*, addressing risks only after they have been identified and shown to be harmful. This stems from a traditional suspicion of government regulation in U.S. law, which requires an extensive factual record demonstrating “significant risks” to justify regulatory action aimed at protecting public health from environmental contaminants. This foundational norm of U.S. legal culture makes precautionary regulation more difficult, as the government must first assemble substantial evidence to support its actions. In this sense see G. Charnley – E. D. Elliott, *Risk versus Precaution: Environmental Law and Public Health Protection*, in *Envtl. L. Rep. News & Analysis*, 32, 2002, 1036 ff and O. Renn – E. D. Elliott, *Precautionary Regulation of Chemicals in the US and EU*, in J. Hammit – M. Rogers – P. Sand – J. B. Wiener (eds.), *The reality of Precaution. Comparing risk regulation in the United States and Europe*, 2013, ff. 224. Furthermore, for a critique on the European approach see the seminal work of G. Majone, *The Precautionary Principle and Its Policy Implications*, in *Journal of Common Market Studies*, 40, 2002, 89 ff. Majone critically examines the EC's promotion of the precautionary principle, arguing that while it has a role in addressing imminent and irreversible risks, its broader application is fraught with logical, economic, and legal shortcomings. He warns that it can distort regulatory priorities, enable protectionism, and undermine international cooperation, highlighting a sharp contrast to the U.S. risk-based, evidence-driven regulatory framework.

²⁸ Such as, for instance, the EU Food Safety Regulation. See European Parliament and the Council, Regulation (EC) No 178/2002 of the European Parliament and of the Council of 28 January 2002 laying down the general principles and requirements of food law, establishing the European Food Safety Authority and laying down procedures in matters of food safety, OJ L 31, 1.2.2002, 1 ff, art. 7.

²⁹ K. De Smedt – E. Vos, *The Application of the Precautionary Principle in the EU*, in H. A. Mieg (eds.), *The Responsibility of Science. Studies in History and Philosophy of Science*, 2022, 163 ff.

³⁰ Zander explains that in international law the precautionary principle is often used based on scientific discovery. However, the author points out, that international practice has also permitted the use of the precautionary principle «when the scientific evidence is so scarce that it is not possible to carry out a meaningful assessment of it». J. Zander, *The Application of the Precautionary Principle in Practice. Comparative Dimensions.*, 2010, 1 ff, 73. For an account that does not fully adhere to a sharp distinction between risk-based regulation and the precautionary principle, conceiving the latter as a technique of anticipatory risk governance, see F. De Leonardis, *Il principio di precauzione nell'amministrazione di rischio*, Milano, 2005, 1 ff.

³¹ This approach is coherent with the overall European regulatory approach in the technological domain. Indeed, the EU has longstanding reliance on risk-based regulation in various sectors, which has allowed policymakers to strike a balance between innovation and precaution. Risk-based regulation itself is a flexible tool, capable of taking different forms depending on how much of the precautionary principle is incorporated. More recently, this approach has been central to the EU's strategy for regulating the digital

More specifically, the notion of risk is used to offset the horizontal regulatory approach, and – in the intention of the legislator – prevent posing too high a burden on innovation³², with the concept of risk broadly defined with reference to health, safety and fundamental rights³³. However, reference to risk represents primarily a narrative used to justify regulatory intervention, rather than the specific assessment of a well-defined concern or potential threat. Said otherwise, the very classification pursuant to different levels of risk is nothing more than a pure conventional and synthetic expression, that does not necessarily correspond to an objective, much less measurable, truth. Indeed, no criterion nor methodology are identified or described by the EC that allows to determine why a given application, and/or class of applications ought to be deemed of greater or lesser concern. One might justifiably wonder why manipulation (art. 5, para. 1, lit. a) AIA) should warrant outright prohibition³⁴, while retrospective remote biometric identification systems are merely classified as potentially high-risk, and deepfakes³⁵ are neither prohibited nor even

domain. Beginning with the Digital Single Market Strategy, risk-based regulation has been applied (albeit with varying approaches) across diverse areas such as data protection, online content and platform governance, cybersecurity, and the regulation of digital products and services. The common goal of this regulatory technique is to mitigate risks proactively by guiding technological development and usage as they emerge, rather than merely responding to harms after they occur. In the context of emerging technologies, such as AI, this strategy operates on the normative assumption that the benefits of technological adoption outweigh the potential risks, provided that these risks are effectively managed. In this sense, see M. E. Kaminski, *The Developing Law of AI: A Turn to Risk Regulation*, in *University of Colorado Law Legal Studies Research Paper No. 24-5*. For a more general account of European risk-based regulation in the digital age, and of the multifaced approaches it adopts, see G. De Gregorio – P. Dunn, *The European Risk-Based Approaches: Connecting Constitutional Dots in the Digital Age*, in *Common Market Law Review*, 52, 2022, 473 ff and for an account of the risk-based approach in the AIA see E. Longo, *La disciplina del “rischio digitale”*, in F. Pizzetti – S. Calzolaio – A. Iannuzzi – E. Longo – M. Orofino (eds.), *La regolazione europea della società digitale*, Torino, 2024, 53 ff, 72.

³² As Mahler explains, the risk-tiered approach has provided a limitation on the obligations imposed upon non-problematic applications. See T. Mahler, *Between risk management and proportionality: The risk-based approach in the EU’s Artificial Intelligence Act Proposal*, in *Nordic Yearbook of Law and Informatics*, 2022, 246 ff.

³³ N. A. Smuha – K. Yeung, *The European Union’s AI Act: beyond motherhood and apple pie?*, in N. A. Smuha (eds.), *The Cambridge Handbook on the Law, Ethics and Policy of Artificial Intelligence*, Cambridge, 2025, 236 ff.

³⁴ The increasing integration of AI into interpersonal and digital interactions poses serious risks, especially regarding the manipulation of human emotions and trust. Generative AI technologies are increasingly being exploited to commit deception and fraud on a massive scale. One example is the rise of AI-enhanced romance scams, as discussed in the 2024 Alan Turing Institute Report by S. Moseley, *Automating Deception: AI’s Evolving Role in Romance Fraud*, 2025, 7 ff, which outlines how LLMs can create convincing fraudulent narratives and synthetic identities. These technologies empower not only individual scammers but also facilitate industrial-scale criminal operations, producing fabricated personas and backstories that evade conventional detection systems. Furthermore, AI-driven psychological profiling allows fraudsters to identify and exploit individual vulnerabilities with unparalleled efficiency, intensifying the impact of these manipulative practices and increasing human vulnerability to digitally enabled threats.

³⁵ Despite the increasing prevalence and harm of deepfake technologies, particularly non-consensual sexual deepfakes, current EU legislation fails to provide an adequate

considered high-risk.

While in certain cases the solution reached by the European legislator is commendable – e.g. prohibiting manipulatory AISs³⁶ (despite the limits imposed by the norm may still be less than satisfactory) – there appears to be no objective grounds or methodologies that account for the conclusion reached. There is no ranking of the fundamental rights potentially affected, that would justify such conclusions, nor a description of a balancing technique between the very same rights being at once positively and negatively impinged upon by an identical application.

In such a perspective, the very notion of risk regulation seems to function as a mere synthetic label used to justify a political choice, and its primary purpose is that of offsetting the unacceptable effects of a horizontal approach that attempts to encompass all existing and foreseeable technological applications. This attempt to reduce all technology to a “one size fits all” regulatory framework has been heavily criticized, as it leads to the «[...] substitution of technical complexity with legal complexity, which does not always lead to more straightforward regulation overall»³⁷.

2.3. The Horizontal Approach and the Importance of Definitions

The coupling of a risk-based categorization with a horizontal regulatory approach was conceived as a solution to the challenge of addressing heterogeneous AISs that can differ significantly from one another, in both function and technical design.

Indeed, the extremely broad, all-encompassing, definition of AIS put forth by art. 3 AIA is a choice precisely intended to ensure maximum

regulatory framework. Not only the AIA, but also the Digital Services Act (DSA) fall short of directly addressing the specific and severe risks posed by such content. Notably, the AIA, despite its human-centric rhetoric, classifies deepfake generation merely as a “limited risk” activity, subject only to minimal transparency obligations, and omits any explicit prohibition of non-consensual sexual deepfakes. This reveals a substantial regulatory gap; as a consequence, serious and harmful applications of AI, such as the creation of non-consensual sexual deepfakes, remain effectively unaddressed within the EU’s emerging AI governance framework.

³⁶ There are many domains in which the effects of a given technology are inherently ambiguous. For instance, PARO (an advanced interactive therapeutic robot designed to support patients with dementia, Alzheimer’s and other cognitive disorders) is generally considered unproblematic, largely due to its simplicity and long-standing presence. However, despite its therapeutic benefits, it can still mislead users by simulating emotional responses, thereby raising ethical concerns about deception and informed interaction, particularly in vulnerable populations. See A. Bertolini, *Human-Robot Interaction and Deception*, in *Osservatorio del diritto civile e commerciale, Rivista semestrale*, 2018, 131 ff; A. Bertolini, *The subtle line between personalization and user manipulation in a European regulatory perspective. A proposal for a technology-assessment methodology for Artificial Intelligence Systems*, in *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2024, 4479 ff.

³⁷ M. Almada, *Two Dogmas of Technology-neutral Regulation*, in *ssrn.com*. On this point, see also A. Bertolini, *Artificial Intelligence does not exist! Defying the technology-neutrality narrative in the regulation of civil liability for advanced technologies*, in *Europa e diritto privato*, 2022, 369 ff.

technological neutrality³⁸. More so, the AIA does not regulate AI technology as such, but rather its use.

That “technology neutrality” is in fact one of the foundational principles of European AI regulation³⁹, as clearly state by the Commission⁴⁰.

However, the very notion of technological neutrality is ambiguous and could lead to contradictory interpretations, as an identical legal provision may be deemed neutral or not according to the chosen perspective. Indeed, a functional understanding requires that legislation should not treat equivalent technologies (those capable of serving similar functions in the same context) in different ways, thereby upholding the principle of non-discrimination⁴¹. A substantive notion, instead, entails that the rationales that a given piece of legislation pursue should not vary across domains; in such a perspective dedicated norms may be required, differentiating between technologies⁴² to ensure an identical treatment – and/or user protection – is granted irrespective of the system used⁴³. The two perspectives could

³⁸ The original definition of AIS in the AIA referred to a specific set of programming techniques, and was thence criticized for lacking sufficient technological neutrality. This approach risked both obsolescence, by failing to capture future AI development, and over-inclusion, by potentially bringing under regulation certain systems not intended to be covered. As AI technologies evolve rapidly, new methodologies may emerge that fall outside the scope of Annex I, thereby escaping regulation despite posing comparable risks. Conversely, some traditional software systems (particularly those using statistical or logic-based operations) might have been unintentionally captured within the regulatory perimeter, despite lacking the autonomous or adaptive capabilities typically associated with AI. In this sense, see M. Almada, *Delegating the Law of Artificial Intelligence. A Procedural Account of Technology-Neutral Regulation*, European University Institute, 2024; 1 ff, 31.

³⁹ *Ibid.*, 73.

⁴⁰ «The proposal [of AIA] sets a robust and flexible legal framework. On the one hand, it is comprehensive and future-proof in its fundamental regulatory choices, including the principle-based requirements that AISs should comply with. On the other hand, it puts in place a proportionate regulatory system centred on a well-defined risk-based regulatory approach that does not create unnecessary restrictions to trade, whereby legal intervention is tailored to those concrete situations where there is a justified cause for concern or where such concern can reasonably be anticipated in the near future. At the same time, the legal framework includes flexible mechanisms that enable it to be dynamically adapted as the technology evolves and new concerning situations emerge» in this sense, see European Commission, Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain union legislative acts. COM/2021/206 final, Brussels, 2021. See also art. 12, 14, 15 and 22 GDPR and, on the right to have an explanation, *ex multis*, S. Wachter, *Limitations and Loopholes in the EU AI Act and AI Liability Directives: What This Means for the European Union, the United States, and Beyond*, in *Yale Journal of Law & Technology*, 26, 2024, S. Wachter – Brent Mittelstadt – F. Luciano, *Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation*, in *International Data Privacy Law*, 7, 2017, 76 ff.

⁴¹ C. J. Craig, *The Relational Robot: A Normative Lens for AI Legal Neutrality—Commentary on Ryan Abbott, The Reasonable Robot*, in *Jrslm Rev Legal Stud*, 25, 2022, 24 ff, 35.

⁴² *Ibid.*, 37.

⁴³ An example is that of Recital 15 GDPR which ensures that users may obtain an explanation about a given decision, regardless of the technology used to automate it, explicitly discriminating against opaque AI solutions, see M. Almada, *Delegating*, cit., 30; M. Hildebrandt – L. Tielemans, *Data Protection by Design and Technology Neutral Law*, in *Computer Law & Security Review*, 29, 2013, 509 ff.

clearly be at odds⁴⁴.

The formulation of the AIA appears at least twofold. On the one hand, it seeks to uphold a principle of non-discrimination between technologies⁴⁵, which is why the regulatory approach focuses on uses rather than on specific tools⁴⁶. On the other hand, it aims to establish a future-proof system⁴⁷, capable of being relevant despite rapid technological innovation. The latter is characteristic of legal systems, which can adapt, primarily through interpretation, once clear principles are established. However, this flexibility comes with compromises in terms of ex ante legal certainty, foreseeability, and consistency of outcomes, all of which affect market uniformity, particularly across MS (*infra*)⁴⁸.

On closer examination, the first perspective appears neither particularly valuable nor coherently implemented in the AIA. Treating all technologies equally is not desirable, as some better promote fundamental rights or support specific policy objectives⁴⁹, and therefore should be favoured.

In such sense the AIA is also not fully neutral, providing relevant discrimination between technologies, also through the so-called risk categorization. The latter, in fact, is not truly risk-based (see §2.2) as no criteria are

⁴⁴ An example shows the practical implications of distinguishing between functional and substantive neutrality. Scholars have debated whether individuals have the right to an explanation of automated decisions involving their personal data. If AISs are used, technical tools like explainable AI are needed. This meets GDPR Recital 15's idea of neutrality by allowing explanations regardless of the technology. Cfr. M. Almada, *Delegating*, cit., 30.

⁴⁵ See D. Kwak, *No More Strategic Neutrality on Technological Neutrality: Technological Neutrality as a Bridge Between the Analogue Trading Regime and Digital Trade*, in *World Trade Review*, 2022, 18 ff; W. J. Maxwell – M. Bourreau, *Technology neutrality in Internet, telecoms and data protection regulation*, in *Computer and Telecommunications Law Review*, 2015, 1 ff.

⁴⁶ European Commission, Proposal for a Regulation COM/2021/206 final, cit., 12 ff describes the definition of AI as technology neutral defining it as «software that is developed with one or more of the techniques and approaches listed in Annex I and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with...».

⁴⁷ *Ibid*, 12 ff mentions that «[t]he definition of AI system in the legal framework aims to be as technology neutral and future proof as possible».

⁴⁸ That of ensuring a non-fragmented market is indeed one of the motivating factors for the adoption of a uniform European regulation in the field of AI.

⁴⁹ This is the case, for instance, of robotic prostheses that can enhance the rights of people with disabilities. See A. Bertolini, *Robotic prostheses as products enhancing the rights of people with disabilities. Reconsidering the structure of liability rules*, in *International Review of Law, Computers & Technology*, 29, 2015, 116 ff. This stance is indeed recognized also at the international level: for instance, art. 4 of the United Nations Convention on the Rights of Persons with Disabilities (A/RES/61/106) includes, among the general obligations of States, the duty «to undertake or promote research and development of, and to promote the availability and use of new technologies, including information and communications technologies, mobility aids, devices and assistive technologies, suitable for persons with disabilities, giving priority to technologies at an affordable cost». This provision underscores that new technological advancement in enhancing the autonomy, inclusion, and overall quality of life of persons with disabilities are particularly welcomed and can nowadays become a key instrument to close existing gaps in access and opportunity, particularly for the most vulnerable groups in society.

provided to enable an objective assessment of risk. Its categorisation thus does not follow from necessity but reflects a specific policy choice that is, by definition, not neutral. When extremely different applications – such as a lower limb exoskeleton⁵⁰ and an algorithm used for decisions affecting employment relationships (such as termination of labour contracts)⁵¹ – end up classified under the same risk level (e.g. high-risk) it seems a questionable outcome, in particular considering the profoundly different technological characteristics and areas of application of each, and subsequently the risks they may give rise to.

At the same time – and as a consequence of problematic and ambiguous definitions such as the one of art. 6 AIA (see §4.2-4.5 and §6) – discrimination may arise between similar applications that differ only in limited technological aspects, and thus may fall either within or outside the high-risk category. Given the complexity of the definition and the numerous factors to be assessed for each AIS, it is also plausible that not all disparities between otherwise comparable technologies are intentional, but rather the result of the room left to interpretation.

As per the possibility of the overall system to preserve its relevance *vis a vis* future innovation because of its horizontal, all-encompassing, and thence neutral, regulatory approach, it is also questionable. More specifically, the framework's future adaptability is not consistently ensured⁵², and overly broad constraints that fail to account for technological specificities may, in fact, stifle innovation⁵³. The AIA is indicative of the uncertainty and short-circuits that technology neutrality engenders when it lacks precise and robust foundations.

Indeed, while the intent is to prevent regulatory bias, neutrality may give rise to interpretative ambiguities, making consistent and thus foreseeable enforcement difficult. This, in turn, may disproportionately benefit established players, who can navigate vague regulations – and absorb the costs

⁵⁰ An exoskeleton would be considered a h-AIS according to art. 6, para. 1, which in turn refers to Annex I. In the case of exoskeletons the parameters to be observed would be those set out by the Medical Devices Regulation (EU) 2017/745) and/or the Machinery Directive (2006/42/EC), depending on its application.

⁵¹ Algorithms use in employment relations are considered to be h-AIS according to art. 6, para. 2, which refers to the list in Annex III. Specifically point (4) of the Annex specifies that an AIS will be high-risk if applied to employment, workers management and access to self-employment

⁵² See B. A. Greenberg, *Rethinking Technology Neutrality*, in *Minnesota Law Review*, 100, 2016, 1495 ff, 1524, where the author argues that technology neutrality faces a fundamental problem of prediction: legislators cannot adequately foresee how laws should apply to emerging technologies until those technologies are known. As a result, laws intended to be future-proof are often poorly tailored, having been implicitly designed with existing technologies in mind. This predictive limitation might undermine the goal of adapting legal frameworks to paradigm-shifting innovation.

⁵³ *Ibid*, 1543, Greenberg critiques the illusion of neutrality in lawmaking, arguing that decisions about which technologies are subject to neutral treatment inevitably reflect social and political value judgments. Attempts to pre-emptively include future technologies may inadvertently create discrimination by failing to anticipate novel use cases. He advocates for a model of technological discrimination that combines neutrality with specificity, enabling more tailored, adaptive, and innovation-friendly regulation while enhancing legal clarity and policy alignment.

of *ex ante* compliance as well as the risks of potential *ex post* litigation – far more effectively than smaller innovators⁵⁴.

Ultimately, a persuasive account of technological neutrality, that perfectly fits the narrative embedded in the AIA, is that which considers neutrality as a technique to delegate choices to other actors. Indeed, the predominantly top-down nature of the AIA's risk categorization can be highly problematic. The Act itself, together with the EC's power under art. 7 to amend the Annexes, outlines the risk categories and the characteristics that place AISs into each category.

Although this structure is intended to create consistency, it may give rise to ambiguity, stemming primarily from (i) the breadth and fluidity of the defining criteria and (ii) the discretion policymakers retain in determining which systems qualify as high-risks or others⁵⁵. Such a top-down approach can be arbitrary: a system's "risk level" may hinge on interpretive nuances, which can shift if the Commission revises the Annexes to reflect new priorities (art. 7 AIA)⁵⁶.

Moreover, unclear or evolving definitions have tangible consequences. A product labelled "high-risk" must meet heightened compliance requirements while a borderline system that avoids high-risk classification might sidestep critical safeguards. The net result is regulatory uncertainty on both ends: providers face unpredictable obligations, and regulators risk missing truly dangerous applications that do not neatly fit the AIA's categories.

Further complexity arises from the partial flexibility granted to providers of h-AIS. As discussed in the following section, art. 6 of the AIA allows

⁵⁴ W. J. Maxwell – M. Bourreau, *Technology neutrality in Internet*, cit., 2.

⁵⁵ In particular, given the concentration of regulatory and administrative powers in the hands of a single entity (the EC) and in the absence of an independent authority tasked with overseeing the matter. In this sense, see, *ex multis*, S. Calzolaio, *Autorità, governo, attuazione dell'AI Act*, in F. Pizzetti – S. Calzolaio – A. Iannuzzi – E. Longo – M. Orofino (eds.), *La regolazione europea dell'intelligenza artificiale nella società digitale*, Torino, 2024, 139 ff.

⁵⁶ The power granted to the EC to modify the lists of high-risk systems recalls legislative techniques already present in certain national legal systems, such as the so-called "*norme penali in bianco*" or "*Blankettstrafgesetz*". These are criminal provisions that defer the definition of their prescriptive content to external sources, often of secondary or administrative rank, thus delegating to these sources the concrete determination of criminally relevant conduct. This legislative technique has been widely criticized in legal scholarship, as it entails an extremely broad delegation of normative power, potentially lacking clearly predetermined limits, and potentially undermines the principle of legality, particularly in its aspects of statutory reserve and definiteness. In this sense see, among others, V. Manes, *L'eterointegrazione della fattispecie penale mediante fonti subordinate, tra riserva "politica" e specificazione "tecnica"*, in *Rivista italiana di diritto e procedura penale*, 2010, 84 ff. Similarly, the power attributed to the EC risks resulting in a substantial expansion of its decision-making authority, especially with respect to determining which applications may qualify as high risk and what does not. Such power is exercised outside the bounds of democratic oversight, bypassing both democratic debate and the ordinary legislative process. The criticisms traditionally directed at "*norme penali in bianco*" might therefore also be relevant, *mutatis mutandis*, to the power of the EC, insofar as both raise concerns regarding democratic legitimacy and the protection of fundamental guarantees for those subject to the norm. Indeed, as further detailed in paragraph 3 of this Section, the Commission's power rests on the questionable assumption that risk assessment is a purely technical exercise. However, such determinations inevitably involve political choices and removing them from open political debate and legislative process might cause significant democratic *vulnus*.

providers listed in Annex III to argue that their systems do not actually pose the designated risk, effectively introducing a bottom-up mechanism (art. 6, para. 4, AIA). While this provision could, in principle, refine the top-down approach by limiting the Commission's power and inviting a more context-sensitive analysis, it has in practice greatly amplified existing uncertainty. Providers must conduct costly and complex self-assessments, while authorities maintain the latitude to reject or reinterpret claims based on evolving regulatory expectations. Coupled with the Commission's power to revise the Annexes, this results in a moving target for compliance. In this way, the AIA's risk-based classification becomes a conceptual junction, encompassing a broad spectrum of technical systems under one legislative umbrella while simultaneously allowing for case-by-case exceptions. Rather than offering clarity, the interplay of broad definitions, top-down categories, and partial provider-initiated reclassification create a confusing regulatory environment. Stakeholders struggle to predict their compliance path, and authorities grapple with mitigating risks to health, safety and fundamental rights effectively, as it will be further address in the following Section.

3. The Midday – Article 6 and the Definition of h-AIS: Annex III and the High-risk Categories.

3.1. The Reference to Annex III

Art. 6 of the AIA identifies the high-risk applications by reference to Annex III (see art. 6, para. 2, AIA) or by imposing two criteria that must be met together (para. 1).

The reference to Annex III is particularly significant, as all AISs listed therein are automatically classified as h-AIS⁵⁷, and should thence precisely determine what applications are subject to the most stringent regulation contained in the AIA. However, for each of these eight areas, Annex III identifies specific conditions, objectives or purposes that the AIS needs to respectively meet or pursue to be deemed an h- AIS. The stated rationale is that all those applications negatively impact on the safety, health, or fundamental rights of individuals.

However, the rigid categorization of Annex III raises several concerns. First, it remains unclear what empirical evidence or criteria were used to define these risk categories and compile the list of covered areas and applications. Second, and more fundamentally, the existence of these risk categories demonstrates that the AIA departs from any claim of being truly “technology neutral”. By selecting which AI applications warrant stricter oversight, the EU legislator made normative and political decisions about which uses are more problematic, rather than abiding by a purely

⁵⁷ Annex III outlines eight areas of application: 1) Biometrics; 2) Critical infrastructure; 3) Education and vocational training; 4) Employment, worker management, and access to self-employment; 5) Access to and enjoyment of essential private services and public services/benefits; 6) Law enforcement; 7) Migration, asylum, and border control management; 8) Administration of justice and democratic processes.

objective or “value-free” process. It is certainly acceptable, and often necessary, for some applications to be subject to stricter regulations, and it is inherently difficult to regulate, under uniform provisions, applications that differ fundamentally. However, in the AIA, the criterion used to assimilate and distinguish these applications is not grounded in a precise or empirically substantiated measurement of risk. Rather, what is referred to as “risk” functions more as a summary expression of a political and normative judgement about potential danger. This kind of judgement, while legitimate, is neither entirely objective nor based on a consistent, transparent evaluation of harm or likelihood. Indeed, no scheme of balancing is either proposed or framed: the classification of AISs into risk levels reflects a political choice rooted in specific priorities and societal values, that however are not narrowly nor clearly defined. For instance, this can be illustrated by the regulation of AIS that involve emotion recognition, particularly in workplace or commercial contexts. While such systems may significantly impact individuals’ privacy, their classification as high-risk is not accompanied by an explanation of how these concerns are weighed against other competing interests, such as, for instance, economic freedom or business innovation. The AIA does not articulate the rationale behind prioritizing the right to privacy over companies’ freedom to develop and deploy technologies that could, for example, increase productivity or enhance customer service. This reflects a conscious judgement that certain technologies or their uses require heightened scrutiny, while others are relatively less worrisome. However, as noted above, this judgement is not – and cannot be – explicitly motivated and/or objectively defensible.

The risk categorization process, therefore, cannot be purely neutral because it necessarily involves subjective assessments about which harms or rights matter most, a question that depends on shifting legislative, ethical and cultural standpoints. Hence, the AIA’s framework embodies political choices that stand in tension with the notion of technology neutrality, revealing that regulators do, in fact, treat distinct AI uses differently in line with societal objectives. Far from being undesirable, these distinctions are often necessary to protect public welfare and fundamental rights, but they underscore why the idea of a fully neutral approach to AI technology is untenable, both in principle and in practice.

In addition to that, and as further evidence of what stated above, many AI applications are considered high-risk solely because they operate within one of these sectors, even if they do not necessarily pose a significant risk of harm. For instance, AISs used for monitoring and detecting prohibited student behaviour during exams in educational and vocational training contexts are automatically categorized as high-risk, irrespective of their technical specifications and of their actual potential to infringe fundamental rights, safety or health⁵⁸. By contrast, certain AI tools deployed in the

⁵⁸ On the challenges posed by AI in educational settings, particularly regarding the balance between academic integrity and individual rights, see J. A. Oravec, *AI, biometric analysis, and emerging cheating detection systems: The engineering of academic integrity?*, in *Education Policy Analysis Archives*, 30, 2022, 1 ff and D. R. E. Cotton – P. A. Cotton – S. J. Reuben, *Chatting and cheating: Ensuring academic integrity in the era of ChatGPT*, in *Innovations in Education and Teaching International*, 61, 2024, 228 ff. The authors highlighted how surveillance-

healthcare sector (such as mental health chatbots that interact with users without human oversight) may not automatically fall into the high-risk category, even though their impact can be substantial.

Perhaps for these reasons, art. 6 also introduces exceptions to the automatic classification of all AI applications mentioned in the sectors listed as high-risk. However, these exceptions and counter-exceptions do not clarify the legal framework; rather they complicate it further.

3.2. Exceptions and Future Amendments

The first exception is introduced in art. 6, para. 3, which provides that AISs listed in Annex III may not be considered as high-risk when they do not pose a significant risk of harm to the health, safety or fundamental rights of natural persons. According to the same provision, this applies when at least one of the four criteria outlined is met: when the AIS performs a simple procedural task, supports or refines a human decision that has already been made, identifies patterns without influencing decisions, or merely prepares input for a later human determination (lit. a), b), c) and d))⁵⁹. In such cases, the system, although formally falling under one of the categories in Annex III, should not be classified as high-risk.

This exemption clearly reflects an effort to avoid overregulation by excluding borderline use cases that do not materially influence decision-making outcomes. The underlying rationale is to prevent disproportionate compliance burdens for applications that, while formally included in high-risk areas, play only a limited or supporting role in practice. However, the provision relies on vague and open-ended language, such as “not materially influencing” or “narrow procedural task,” which makes it difficult to determine *ex ante* whether a specific system qualifies for the exemption⁶⁰.

based cheating detection technologies and tools like ChatGPT raise unresolved tensions between privacy, fairness, and institutional control, often without a clear framework for balancing these competing values.

⁵⁹ According to art. 6, para 6 of the AIA, an AIS referred to in Annex 3 should not be considered to be high-risk where any of the following conditions is fulfilled: «(a) the AI system is intended to perform a narrow procedural task; (b) the AI system is intended to improve the result of a previously completed human activity; (c) the AI system is intended to detect decision-making patterns or deviations from prior decision-making patterns and is not meant to replace or influence the previously completed human assessment, without proper human review; or (d) the AI system is intended to perform a preparatory task to an assessment relevant for the purposes of the use cases listed in Annex III.»

⁶⁰ Scholars have already pointed out that these exceptions can be problematic and seems to add complexity in the AIA's regulatory architecture causing further legal uncertainty. See, for instance, N. A. Smuha, *Nathalie A Smuha, Algorithmic Rule by Law: How Algorithmic Regulation in the Public Sector Erodes the Rule of Law*, 2024, 1 ff, 105; M. Ebers, *Truly Risk-Based Regulation of Artificial Intelligence: How to Implement the EU's AI Act*, in *Stanford-Vienna European Union Law Working Paper No. 101*, 2024, 1 ff and S. Wachter, *Limitations and Loopholes in the EU AI Act and AI Liability Directives: What This Means for the European Union, the United States, and Beyond*, cit., 671 ff. More in general, for a comprehensive overview of the various provisions of the AIA, along with a detailed analysis of the obligations set forth by the Act, see P. Voigt – N. Hullen, *The EU AI Act. Answers to Frequently Asked Questions*, 2024, 1 ff.

The practical implications of the four criteria set out in art. 6, para. 3, are significant and may give rise to several key issues concerning its application, potentially leading to regulatory fragmentation across MS and creating legal uncertainty for providers and deployers.

First, determining when an AIS influences decision-making or is genuinely intended only to detect decision-making patterns, without replacing or unduly influencing prior human assessments, is extremely difficult. For example, in the context of a judicial decision-support tool, it may be difficult to distinguish between an AI that merely highlights trends in sentencing and one that subtly guides judges towards a particular outcome. One of the central challenges lies in defining how an AIS “merely detects” patterns rather than shaping or overriding the human decision-making process. Even when an AIS is designed to provide limited assistance, the way its outputs are generated and presented can significantly affect users’ behaviour. In many cases, individuals perceive algorithmic recommendations as objective or expert, which can lead them to rely on these outputs, even if the system is nominally only assisting⁶¹. This illustrates how system explicability and user perception are interlinked. Indeed, a lack of clarity around how the system arrives at its suggestions often results in humans deferring to, rather than merely consulting, the AI’s judgment, striking examples of which can be found, for instance, in the field of medical diagnostics⁶².

Consequently, the use of the system (as shaped by its design and presentation) may substantially impact the final decision, calling into question the assumption that human assessments remain unaffected. This practical trend, exacerbated by the growing number of highly accessible applications of generative AI, underscores the fear that users will increasingly accept and over-rely on AI-generated suggestions as authoritative. From a scientific perspective, recent studies, including a comprehensive survey study by Microsoft Research⁶³, are beginning to empirically document this very phenomenon. These studies show that increased trust in AI correlates with reduced critical thinking and cognitive effort, suggesting that

⁶¹ H.-P. Lee – A. Sarkar – L. Tankelevitch – I. Drosos – S. Rintel – R. Banks – N. Wilson, *The Impact of Generative AI on Critical Thinking: Self-Reported Reductions in Cognitive Effort and Confidence Effects From a Survey of Knowledge Workers*, in *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, 2025, 1 ff.

⁶² See J. Wei – E. Verona – A. Bertolini – G. Mengaldo, *Explainability matters: The effect of liability rules on the healthcare sector*, in *arXiv.org*, 1 ff. The Authors illustrate that even when an AIS is nominally deployed as a mere “assistant”, its perceived objectivity can prompt decision-makers to rely on its recommendations more than intended or warranted. In particular, they note that especially if the AIS functions as an “oracle” (i.e., provides a diagnosis or conclusion with no accompanying explanation), the human user (in the article, the medical practitioner) may strategically conform to the system’s output to minimize potential liability. Indeed, when the reasoning is opaque, individuals often defer to the presumed accuracy or authority of the AIS. The authors underline that this behaviour can lead to a form of defensive medicine whereby users adopt the AIS’s verdict primarily to shield themselves from potential liability, rather than to serve the best interests of the patient.

⁶³ H.-P. Lee – A. Sarkar – L. Tankelevitch – I. Drosos – S. Rintel – R. Banks – N. Wilson, *The Impact of Generative AI on Critical Thinking: Self-Reported Reductions in Cognitive Effort and Confidence Effects From a Survey of Knowledge Workers*, cit., 1 ff.

overreliance is not just anecdotal, but has measurable cognitive and behavioural consequences in real-world knowledge work⁶⁴.

Second, ensuring consistency in how such assessments are conducted across different AISs and sectors remains a complex task. As mentioned in Section I, AISs vary greatly depending on their technological design and intended use. The same high-level criteria may yield different results when applied to distinct domains (for example, healthcare versus human resources) because the risk profile, data inputs, and societal implications of these technologies differ substantially.

This diversity underscores the difficulty of using a generalized or “neutral” approach to classification: while technology neutrality suggests regulating at a high level of abstraction, contextual factors can render the same criteria either too lenient or too strict, thereby resulting in uneven or inconsistent enforcement. For instance, a criterion requiring transparency in decision-making may be relatively easy to satisfy in a recruitment algorithm that filters CVs but could be impractically demanding for a diagnostic AIS that relies on complex, non-interpretable models to identify rare diseases. In the former, a failure to meet transparency standards might lead to minor compliance concerns; in the latter, it could result in denying access to potentially life-saving innovations.

Third, questions may arise regarding the extent and effectiveness of external and independent oversight in monitoring these evaluations. It remains unclear who will be responsible for evaluating and enforcing these provisions, and how such oversight mechanisms will operate in practice. In theory, robust external supervision (such as independent auditing bodies or regulatory agencies) could ensure that providers and deployers accurately classify their AISs under art. 6, para. 3. However, the Act does not explicitly define the extent to which such external oversight will be required or how it should be standardized across Member States. Without well-defined guidelines on independent review, there is a risk that evaluations will be inconsistent or that some providers may understate their system’s influence on decision-making processes to avoid stricter compliance burdens. In addition to these issues, it is important to note that another provision contained in art. 6, para. 3, may create further uncertainty. The last sentence of art. 6, para. 3, specifically comma 3, introduces an exception to the exception by adding a further limitation related to profiling⁶⁵. Even if one of the four conditions that would normally exclude an AIS from being classified as high-risk is met, the exemption does not apply if the system engages in profiling, as defined by the GDPR⁶⁶. While this carve-

⁶⁴ N. Kosmyna – E. Hauptmann – Y. T. Yuan – J. Situ – X.-H. Liao – A. V. Beresnitzky – I. Braunstein – P. Maes, *Your Brain on ChatGPT: Accumulation of Cognitive Debt when Using an AI Assistant for Essay Writing Task*, in *arXiv.org*, 2025, 1 ff.

⁶⁵ «Notwithstanding the first subparagraph, an AI system referred to in Annex III shall always be considered to be high-risk where the AI system performs profiling of natural persons».

⁶⁶ When defining profiling, the AIA (art. 3, para. 52 refers to the definition of profiling provided by art. 4, point (4), of Regulation (EU) 2016/679: «profiling means any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyse or predict aspects concerning that natural person’s performance at work, economic situation, health,

out is intended to preserve safeguards against intrusive or discriminatory practices, its scope is not always easy to define. Seemingly benign applications, such as personalized content feeds or basic consumer segmentation, may fall within the definition of profiling and may therefore be deemed h-AIS, while systems used in public sector contexts to detect “deviations from prior decision-making patterns” in areas like law enforcement or welfare eligibility could indirectly steer decisions without being formally classified as profiling or without triggering the high-risk threshold, since they are framed as “mere” procedural or advisory tools.

These unresolved questions, unfortunately, represent only a fraction of the broader interpretative challenges posed by the structure of art. 6, para. 3.

On the one hand, it would appear to establish a presumption *iuris et de iure* of high-risk with regard to the profiling of natural persons, as defined in art. 4, para. 4, of GDPR⁶⁷.

On the other hand, it maintains the efficacy of comma 1 of para. 3, specifically the exemption based on the expansive and ambiguous formula previously examined, which states the absence of substantial risks to the health, safety, or fundamental rights of the individual. However, the provisions of comma 2 of para. 3 are not referenced, even though they should constitute a specification of the first.

This prompts numerous inquiries into the interpretation of the text. The exclusion of the initial clause, «without prejudice to the first subparagraph», would result in the classification of all systems engaging in profiling as high risk, thereby eliminating the necessity for additional evaluation. This conclusion would be consistent with the English version of the text, which employs the expression “notwithstanding”, and with the German version, which utilizes the term “*ungeachtet*”, both of which can be translated as “by way of derogation”. By derogating from the initial paragraph, the exception to the rule would be neutralized, thereby classifying any user profiling system as high risk.

However, in the Italian version of the text, the exception referred to in comma 1 is expressly reserved, thus equating profiling with all the other figures indicated in Annex III. The literal interpretation of the text does not appear to allow for divergent interpretations. Consequently, it does not seem possible to consider that profiling applications are subject to a special, worse, or more stringent regime.

In order to attribute meaning to the provision in comma 3 of para. 3, it is necessary to establish a parallel with the regulations concerning unfair terms in consumer contracts⁶⁸. In that context, it is not possible to demon-

personal preferences, interests, reliability, behaviour, location or movements».

⁶⁷ This is the interpretation given by the Commission, which, to date, is supported by the warning it issues in the EU AI Act Compliance Checker, a novel instrument of the AI Office. This tool has been crafted to clarify the obligations and requirements of the AIA. Once the self-assessment has begun and you have reached the question of whether the AIS can be considered high risk pursuant to art. 6, para. 2, AIA, you must specify whether derogations pursuant to art. 6, para. 3, AIA apply. At this point it is literally warned to note that «an AI system referred to in Annex III shall always be considered to be high-risk where the AI system performs profiling of natural persons (Article 6 (3) AI Act)». See ai-act-service-desk.ec.europa.eu.

⁶⁸ For the parallelism please refer to A. Bertolini, *La definizione di sistemi di Intelligenza*

strate that clauses included in the so-called “black list” (art. 36, para. 2, of the Italian Consumer Code) are not unfair by providing evidence of individual negotiation, without prejudice to other cases of exclusion. In this case, too, it would be necessary to reconstruct a different regime for profiling applications by way of interpretation, i.e., by emphasizing the lack of reference to comma 2 of the same paragraph. Consequently, it can be deduced that, in the context of systems engaged in profiling activities, the criteria delineated in comma 2 from lit. a) to d), are not applicable. This would be equivalent to saying that profiling justifies classifying the system as high risk unless it can be demonstrated that it does not seriously violate the security, health, or other fundamental rights of the individual. Conversely, the developer of a profiling system could exclude its high-risk nature by incorporating the exception cited in the comma 1, without resorting to the cases enumerated in comma 2.

While this interpretation appears more consistent in terms of the formal and literal interpretation of the text, it is not at all satisfactory from a functional and remedial point of view. This raises the question of the very rationality of the provision and its real meaning and scope. Conversely, if the European legislator’s intent is to be considered as having been otherwise, it would be necessary to conclude that a significant translation error has been made. This would be analogous to the situation that transpired when the regulations concerning unfair terms were initially introduced⁶⁹. Given the aforementioned points, it is evident that these considerations collectively contribute to the overall ambiguity of the provision and intensify uncertainty regarding its practical application.

3.3. Amendments to Annex III

In addition to the above, it should be stressed that the list of h-AISs established by the AIA is subject to ongoing updates, as delegated to the EC under art. 7 of the AIA. This delegation grants the Commission the authority to modify the Annex over time, either by adding or removing specific applications or by adjusting the criteria required for systems to qualify as h-AIS. As a result, the regulatory landscape for AI applications is inherently dynamic, with significant implications for providers. However, as already outlined above (§2.3), this power is extremely broad and rests on the fallacious assumption that the definition of risk is the outcome of a purely technical assessment. In reality, such determinations are inherently political, and by removing them from open political debate, the mechanism introduces a potentially significant democratic vulnerability⁷⁰.

Artificiale, in G. Cerrina Feroni – S. Orlando – G. Vettori (eds.), *Il Regolamento europeo sull’intelligenza artificiale. Problemi e Sistema*, Bologna, 2025, 183 ff.

⁶⁹ *Ibid*, 197.

⁷⁰ For an analysis of the vulnerabilities inherent in the digital society and their constitutional implications within the emerging datafied legal order, see, *ex multis*, S. Calzolaio, *Vulnerabilità della società digitale e ordinamento costituzionale dei dati*, in *Rivista italiana di informatica e diritto*, 2, 2023, 14 ff.

For instance, applications initially designed to comply with the requirements for h-AISs might later find themselves subject to a more lenient regulatory regime if changes to the Annex or criteria occur. Conversely, applications not originally classified as high-risk might, at a later stage, be added to the Annex if regulators identify emerging risks associated with their deployment or functionality. This adaptability ensures that the regulatory framework can respond to technological advancements and unforeseen developments, but it also introduces a degree of uncertainty for AI providers. For example, a novel application that was not initially anticipated as problematic during its conceptualization and early design stages might later be deemed a source of significant concern. If this occurs, providers may be compelled to revisit and modify the design of their application to ensure adherence to the newly imposed requirements for h-AISs⁷¹. Such retrospective compliance obligations could be particularly burdensome for applications with extended research, development, and testing timelines. In the field of robotics, for instance, the development cycle for complex applications often spans several years, encompassing stages from initial research and prototyping to rigorous testing and deployment⁷². This extended timeframe increases the likelihood that the regulatory landscape may evolve during the course of development. Consequently, providers may find themselves needing to adapt to new compliance requirements mid-project, which can lead to significant delays, increased costs, and resource reallocation. This is especially challenging for those working on sophisticated, embodied AISs intended for critical or high-stakes environments, where safety and reliability are crucial.

3.4. The “Self-Declaration”

Finally, art. 6, para. 4, allows any provider to conclude that their AIS is not high-risk, even if it falls within the scope of Annex III, provided they document their assessment, register the product and, upon request of national authorities, provide the documentation upon which the assessment is grounded.

This means that a single manufacturer could argue that their device, service, or application should not be classified as high-risk, despite it theoretically falling into that category. To support this claim, manufacturers must document their assessment before placing the system on the market or

⁷¹ See on this E. Commission, *Study on the functions and effects of European standards and standardisation in the EU and EFTA Member States and Annexes*, 2021. This document notes that, particularly in fast-moving technology sectors, European standards are often perceived as lagging behind technological developments. Stakeholders report long intervals between the technical finalization and publication of standards, making them less effective for companies needing to adapt rapidly to new regulatory or market requirements (p. 10; see also pp. 158-159). This temporal misalignment reinforces the concern that developers may face significant redesign burdens if compliance conditions shift mid-development.

⁷² OECD, *Artificial Intelligence in Science: Challenges, Opportunities and the Future of Research*, Paris, 2023; N. Alistair, *Artificial intelligence in science: challenges, opportunities and the future of research*, 2024.

putting it into service and comply with the registration obligation set out in art. 49, para. 2. Only if the national authority requests it, the manufacturer is required to provide the documentation for review.

However, a series of critical ambiguity arises regarding the scope of application of this provision. It remains unclear whether this provision is intended to apply only to systems referred to in Annex III that shall not be considered to be high-risk because they exhibit some of the characteristics outlined in para. 3 (for instance, because they are intended to perform a narrow procedural task, see para. 5) or whether it can be invoked more broadly on the basis of other considerations, eventually more technical in nature, or, again, based on the assessment of the criteria and condition listed for each of the eight categories by Annex III. Furthermore, one might also question whether art. 6, para. 4, could also extend to AISs falling under para. 1 of art. 6.

This lack of clarity may give rise to several criticalities. First, it may lead to divergent interpretations and practices among providers and national authorities, potentially undermining the uniform application of the regulation and creating legal uncertainty for both regulators and providers. Secondly, the distinct regime under art. 6, para. 4, introduces a criterion and a procedure that may prove insufficiently rational and potentially discriminatory in relation to the scope and the entities covered by art. 6, para. 1. This raises concerns regarding the legal soundness of a potential dual-track system.

Thirdly, the process described in para. 4 lacks any formal decision-making stage or mechanism of official endorsement. There appears to be no principle of tacit approval through administrative silence – *i.e.*, no rule whereby an individual's request is automatically approved if the administration fails to respond within a set period⁷³ – nor any systematic evaluation procedure conducted by an independent authority. As a result, the provider is left without definitive confirmation that the classification of the AI system as non-high-risk is correct; and, conversely, even if it were to receive express confirmation from a competent authority, the result would no longer constitute a self-certification but rather a certified declaration. This creates legal uncertainty and raises the risk that the system may ultimately fail to meet the requirements for h-AISs as set out in the regulation.

Even more concerning, this gap may lead to situations where the manufacturer is forced to comply with high-risk obligations only at a later stage (when the system is already fully developed) with potentially serious consequences. This could eventually require a complete or partial redesign of the system, resulting in significant costs and delays.

Without an authoritative validation process, both compliance and accountability are compromised, potentially undermining the effectiveness of the

⁷³ A principle of this kind, for example, exists in Italy under art. 20 of Law No. 241/1990, which states that «in proceedings initiated by an application of a private party for the issuance of administrative measures, the silence of the competent administration is equivalent to acceptance of the request, without the need for further applications or warnings, if the administration does not notify the applicant within the time limit set out in art. 2, para 2 or 3». For an overview, see P. G. Lignani, *Silenzi (diritto amministrativo)*, in *Enciclopedia del diritto*, 1999, 978 ff.

regulatory framework intended to safeguard fundamental rights and public interests. From a procedural perspective, the absence of a clear “review step” means there is no formal moment in which an authority confirms or disputes a provider’s classification. This lack of adjudicative oversight not only deprives the provider of legal certainty but also impedes transparent public accountability, since no authoritative body is required to justify or document its reasoning in approving (or rejecting) the provider’s decision. Notably, from a procedural point of view, formal decision-making processes are seen as essential to ensuring uniform application of regulations and enabling due process, yet here those mechanisms are notably absent. Furthermore, weak *ex ante* oversight can induce moral hazard, since providers may be tempted to classify borderline systems as non-high risk to reduce compliance burdens, all while shifting potential costs (for example harms to users or the public) to a later stage. Without early independent verification, these providers might not internalize the full societal risks of under-compliance. Indeed, rigorous *ex ante* evaluation is often credited with aligning private incentives (profit, speed to market) with broader social welfare goals⁷⁴. If that step is missing, the risk of under-compliance grows, as does the possibility of negative externalities, such as compromised safety or erosion of fundamental rights. In practice, this may erode trust in the market for AISs and weaken the overall regulatory scheme’s efficacy. Consequently, while self-declaration might seem efficient at first glance, it creates a fragile balance between promoting innovation and preserving public safeguards, underscoring the need for a more authoritative and transparent validation mechanism.

In conclusion, considering the latitude left to interpretation by the wording of the norm, there is ample scope for discretionary judgments by different authorities. This ultimately suggests that pursuing such an avenue of self-declaration is feasible but highly risky, both for providers (who lack certainty) and for the individuals, whose fundamental interests the regulation should aim to protect.

4. The Evening – Art. 6, para. 1, of the AIA: Defining h-AISs

4.1. The Complexity of para. 1

Para. 1 defines the other, more general conditions, that lead an AIS to be deemed high-risk, by referencing to pre-existing Union Harmonized Legislation (UHL), listed in Annex I, and formulating two distinct criteria, namely:

- a. «the AI system is intended to be used as a safety component of a product, or the AI system is itself a product, covered by the Union harmonisation legislation listed in Annex I;

⁷⁴ For a general perspective on self-assessment mechanisms, see R. A. Epstein, *The Use and Limits of Self-Valuation Systems*, in *University of Chicago Law Review*, 81, 2014, 109 ff and S. Levmore, *Self-Assessed Valuation Systems for Tort and Other Law*, in *Virginia Law Review*, 68, 1982, 771 ff.

- b. the product whose safety component pursuant to point (a) is the AI system, or the AI system itself as a product, is required to undergo a third-party conformity assessment, with a view to the placing on the market or the putting into service of that product pursuant to the Union harmonisation legislation listed in Annex I».

Both criteria require complex interpretation. On the one hand, the definition of safety component proffered by art. 3, para. 14 AIA is circular in nature and leaves relevant uncertainty about which component is to be deemed relevant for the analysis, especially in complex and interconnected systems (see §§4.2 and 4.3). On the other hand, the wording chosen when referring to third-party conformity—namely it being “required”—, may radically affect the scope of application of the provision (see §§4.4-4.5), causing it to either encompass a broad spectrum of AISs or, instead, significantly limiting its relevance (see §§4.6-5).

4.2. The Safety Component

The first requirement laid out by lit. a) is met both when the AIS overall, or one of its safety components is subject to UHL. However, while the first instance is easy to verify, the latter rests on the ambiguous notion of “safety component” defined by art. 3, para. 14 AIA as a

«component of a product or of an AI system which fulfils a safety function for that product or AI system, or the failure or malfunctioning of which endangers the health and safety of persons or property».

The definition is clearly circular – such is the reference to a “safety function” – and fails to provide clear operational criteria that would allow either a legal or technical expert from objectively concluding that a given subsystem is, in fact, the item upon which the remaining analysis – namely that of lit. b) – ought to be carried out.

Indeed, based on the very wording, it is disputable whether the AI-assisted vision system of a mobile robot operating in a human-populated environment is to be deemed a safety component, or solely its braking system (example 1). While it is certain that a malfunctioning in the sensing and elaboration of the information collected could cause an accident, the same is true most likely for most components in the robots itself. This leads to a potentially unbounded interpretation, in which nearly any subsystem could be relevant simply because its malfunction might contribute, however indirectly, to harm. Such reasoning risks an infinite regress of causal attribution, akin to what arises when the *conditio sine qua non* test is applied without normative constraints: if every factual precondition is treated as equally relevant, the scope of responsibility—and here, by analogy, regulatory attention—expands endlessly⁷⁵.

⁷⁵ Indeed, even the consultation initiated by the EC on June 6th, 2025 identified the notion of safety component as one that could provide a lot of uncertainty to stakeholders. European commission, *Commission Launches Public Consultation on High-Risk AI Systems*, in digital-strategy.ec.europa.eu.

This problem is well known in criminal law’s causality theory, particularly in the context

At the same time, the AIS could be designed in a way that the braking system is always in condition to intervene, overriding all other inputs provided by the AI-controlled components, perhaps on the basis of the inputs provided by independent proximity sensors. It could thence be questioned whether such a technical solution would suffice in excluding the “safety-nature” of the AI-assisted vision system and therefore allow the entire AIS to escape the subsequent qualification as high-risk, and its regulatory implications.

Uncertainty on the side of manufactures could lead to both over- and under-compliance, also depending on the market positioning of the individual player⁷⁶, as well as create an undesirable⁷⁷ potential barrier to entering the market.

Moreover, considering how art. 6, para. 1, AIA itself references to UHL—listed in Annex I—that, in turns, adopts individual definitions of “safety component” for the specific purposes of the domain each directive or regulation insists upon⁷⁸, a clear coordination between those overlapping frameworks appears necessary, and yet it is missing. The Commission has issued a consultation through which stakeholders were called to present their doubts and opinions about some concepts pertaining to AIA. The survey focuses mostly on the definition of h-AIS according to art. 6, para. 2—as discussed in Section “Midday”. Nonetheless, section 1 of the survey opens the space for stakeholders to pose questions related to art. 6, para.

of the *conditio sine qua non* test. Without normative constraints, this test leads to a *regressus ad infinitum*, where any antecedent fact (even remote or incidental ones) could be seen as causally relevant. German legal scholarship addresses this through the doctrine of *objective Zurechnung* which introduces a normative filter to identify legally significant causes. In this sense, see C. G. Roxin, Luis, *Strafrecht Allgemeiner Teil*, 1, 1992. In the Italian legal tradition, a similar concern arises in the distinction between mere physical causation and *rilevanza giuridica*, requiring that the causal link fall within the scope of the protected legal interest. See T. Padovani, *Diritto Penale*, Milano, 2002.

⁷⁶ Indeed, it may be anticipated that a larger company who is less concerned about facing ex-post litigation and sanctions, could take the risk of non-complying under uncertainty, whereas a SME or a startup might prefer to pay a certain, higher ex-ante compliance cost that face potential litigation and sanctions at a later day. As explained by Sommers and Cole, «[b]ecause the costs of avoiding requirements are so high, small firms are left with the choice of complying, at very high costs per dollar of sales, or of choosing partial or non-compliance strategies, hoping that regulatory agencies will fail to detect their non-compliance». P. Sommers – R. J. Cole, *Compliance costs of small and larger businesses*, in *Policy Studies Journal*, 1985, 701 ff. Mendoza et al. also point that firm size directly influence the capability of acquiring information and, thus, compliance: «there is a positive and significant association between firm size and knowledge. Acquiring knowledge of complex regulations may be more difficult for smaller firms». J. P. Mendoza – H. C. Dekker – J. L. Wielhouwer, *Firms' Compliance with Complex Regulations*, in *Law and Human Behavior*, 40, 2016, 721 ff.

⁷⁷ Indeed, «legal standard is uncertain, even actors who behave ‘optimally’ in terms of overall social welfare will face some chance of being held liable because of the unpredictability of the legal rule» J. E. Calfee – R. Craswell, *Some Effects of Uncertainty on Compliance with Legal Standards*, in *Virginia Law Review*, 1984, 965.

⁷⁸ These include Machinery Directive (henceforth MD); Machinery Regulation (MR); Toy Safety Directive (henceforth TSD); Lifts and Safety Components Directive (henceforth LD); Radio Equipment Directive (henceforth RED) and the Medical Devices Regulation (henceforth MDR).

1, and Annex I, specifically “on the concept of a safety component”⁷⁹. In itself the consultation does not provide any guidelines on how to interpret the concept of safety component, nor how to balance possible misalignments between the definition of the AIA and those of other UHL, but it shows that the commission recognizes that there might be problem with the concept of safety component.

At the same time, the lack of certainty arising from the AIA could also be compensated by a systematic interpretation of those concepts, deriving guidance from existing definitions of safety components, as recalled in UHL.

The table below provides a comparative overview of the different notions emerging in the UHL recalled in Annex I. However, please note that most of the UHL referred to in Annex I do not provide a definition of safety component, but refer to concepts that allow the product regulated to comply with the requirements set by that specific directive or regulation. The purpose of the table is also to highlight potential convergences and divergences in the wording and notions technical practitioners are already trained in tackling through their work.

⁷⁹ European commission, *Commission Launches Public Consultation*, cit.

Legislation	Definition of “Safety Component” (or note if none)	Safety function	Independently placed on the market	Failure endangers the safety of persons	Unnecessary for functioning
AI Act	«‘safety component’ means a component of a product or of an AI system which fulfils a safety function for that product or AI system, or the failure or malfunctioning of which endangers the health and safety of persons or property »				
D i r e c t i v e 2006 / 42 / E C (MD), Art 2, lit. c)	«‘Safety component’ means a component: (a) serving a safety function , (b) independently placed on the market , (c) whose failure endangers the safety of persons , and (d) which is not necessary for normal functioning or can be replaced by standard part»				
Regulation (EU) 2023/1230 (MR), Art 3, para. 17	«‘Safety component’ means a physical or digital component, including software, designed to fulfil a safety function , independently placed on the market , whose failure or malfunction endangers the safety of persons , and which is not necessary in order for the machinery or related product to function , or can be replaced with a component that does not fulfil a safety function»				
R e g u l a t i o n 2016/424, (Cableway Installations, CIR)	«safety component’ means any component of equipment or any device intended to be incorporated into a subsystem or a cableway installation for the purpose of ensuring a safety function , the failure of which endangers the safety or health of passengers , operating personnel or third parties»				
D i r e c t i v e 2009 / 48 / E C (TSD)⁸⁰	No definition of “safety component”. Manufacturers are required to protect the health and safety of consumers (Art 4, para. 4, and 10 TSD). Annex II provides a list of specific requirements the manufacturer must observe				
D i r e c t i v e 2014 / 33 / E U (LD)⁸¹	Refers to “safety components for lifts” (art. 8 and Annex III, LD) and provides a list of safety components. Refers also to the application of the MD’s Annex I when it comes to essential health and safety requirements				

⁸⁰ European Parliament and the Council of the European Union, Directive 2009/48/EC of the European Parliament and of the Council of 18 June 2009 on the safety of toys, OJ L 170, 30.6.2009, 1,1.

⁸¹ European Parliament and the Council of European Union, Directive 2014/33/EU of the European Parliament and of the Council of 26 February 2014 on the harmonisation of the laws of the Member States relating to lifts and safety components for lifts, OJ L 96, 29.3.2014, 251.

D i r e c t i v e 2014 / 53 / E U (RED)⁸²	<i>No definition of “safety component.” Establishes essential requirements at product/system level. The RED refers to the safety requirements established by the LVD</i>				
5. Regulation 2017/745 (MDR)⁸³	<i>No definition of the concept of “safety component.” The regulation uses “accessory” and “component” but without legal definition of safety-specific components</i>				

The MD and its successor, the MR represent the exception amongst the UHL listed in Annex I in that they both offer a definition of safety component, based on 4 elements, three of which have what we may consider a technical nature. Indeed, the component needs to have a safety function lit. a), its failure endangers the safety of persons lit. c) but at the same time, it is not necessary for the functioning of the product lit. d). The fourth is, instead, an exception, absent in all other definitions, and refers to independent the placing on the market lit. d), that is to be understood as the possibility that the component is sold onto the market as such. That further reference does not add, however, to the elements that would at least theoretically allow for the identification of those components that indeed play a safety role, and needs not be further commented upon.

The key elements that appear to have inspired the solution enacted by the AIA, as well as that by the CIR, are the reference to both the safety function and the consequence of failure. Those criteria, however, appear again purely circular, incapable of providing any relevant indication. On the one hand, coordination between those definitions appears quite straightforward, in as much as a safety component for the purposes of the AIA will most certainly be considered such also by the MD and MR. The missing reference to the requirement laid down by lit. b) appears, in fact, irrelevant. On the other hand, we may not look at the wording of those definitions to gain a better understanding of what exactly falls under the scope of art. 6, para. 1, lit. a), but maybe the current applications of those, long existing pieces of legislation could indeed provide some guidance.

Radically different approaches are instead adopted by more specific pieces of legislation, such as the LD, that by referring to a very narrowly defined category of homogeneous products – lifts – may replace ample and underdefined concepts with a practical list of components that undoubtedly

⁸² European Parliament and the Council, Directive 2014/53/EU of the European Parliament and of the Council of 16 April 2014 on the harmonisation of the laws of the Member States relating to the making available on the market of radio equipment and repealing Directive 1999/5/EC, OJ L 153, 22.5.2014, 62.

⁸³ European Parliament and the Council, Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC, OJ L 117, 5.5.2017, ibid.

serve a safety function.

In contrast, to legislation that provides precise definitions of safety-related terms, certain sector-specific regulatory frameworks—such as the MDR and the RED—adopt a broader and less prescriptive approach to safety terminology. For example, the MDR establishes, in Annex I, a set of “general safety and performance requirements” that apply to medical devices placed on the EU market (art. 5, para. 2, MDR). These requirements encompass a wide array of considerations related to intended use, risk management, and clinical evaluation of medical devices, rather than defining safety in narrow or technical terms.

Similarly, the RED does not offer a specific definition of a “safety component.” Instead, it explicitly refers to the safety objectives articulated in the Low Voltage Directive (LVD⁸⁴) as sufficiently comprehensive to address the safety aspects of radio equipment. Recital 7 of the RED states that «the objectives with respect to safety requirements laid down in Directive 2014/35/EU are sufficient to cover radio equipment, and should therefore be the reference and made applicable by virtue of this Directive». This cross-referencing illustrates the EU’s regulatory strategy of leveraging existing frameworks to avoid redundancy and promote consistency across sectors.

The LVD itself outlines, in Annex I, general principles of safety, which include protection against hazards originating from the electrical equipment as well as those that may arise due to external influences acting upon the equipment. These principles are intentionally broad, focusing on outcomes rather than prescribing specific technical measures. The use of such generalized language reflects a regulatory emphasis on risk prevention and functional safety outcomes across a wide variety of electrical products.

Particularly relevant is the definition advanced by the MR, both for its greater complexity and sophistication and for the importance of that piece of legislation, applying to possibly the broadest spectrum of products⁸⁵ among all those recalled under Annex I. In particular, the definition formulates an “indifference condition”, whereby the component is intended for safety purposes if it «[...] is not necessary in order for that product to function or for which normal components may be substituted in order for that product to function»⁸⁶. So intended, a safety component is separate from the product, regardless of its necessity for the product’s performance⁸⁷. The concept here is that the component might improve safety, but is not necessary in order for that product to function⁸⁸.

If we applied this notion to the “example 1” above, we could doubt wheth-

⁸⁴ European Parliament and the Council, Directive 2014/35/EU of the European Parliament and of the Council of 26 February 2014 on the harmonisation of the laws of the Member States relating to the making available on the market of electrical equipment designed for use within certain voltage limits, OJ L 96, 29.3.2014.

⁸⁵ Indeed, pursuant to art. 3, para. 17, MR it applies to all machinery defined by art. 3, para. 1 MR.

⁸⁶ European Parliament and the Council, *Regulation (EU) 2023/1230*, cit.

⁸⁷ T. De Graaf – G. Veldt, *The AI Act and Its Impact on Product Safety, Contracts and Liability*, in *European Review of Private Law*, 2022, 803.

⁸⁸ European Parliament and the Council, *Regulation (EU) 2023/1230*, cit.

er the braking system is replaceable, or the AI-assisted vision system. Most likely both could be deemed essential for the functioning – given that the purpose of the machine is to navigate the environment without colliding with obstacles – and thence could possibly not be deemed safety components for the purposes of art. 2, lit. c) MD. Moreover, even if we were to consider the indicative list of Annex V, MD, it does not seem to directly address «example 1». Obviously, the list is indicative and the argument could be made to consider such systems safety components. That would be reinforced by an analysis of Annex II, MR – the successor of the MD – where the indicative list cites (18) Software ensuring safety functions and (19) Safety components with fully or partially self-evolving behaviour using machine learning approaches ensuring safety functions. While this might be an indicative that AIS would fall into the definition of safety component. This doesn't seem to be an improvement, as it still rely on the circular definitions of safety components and safety functions.

In relation to the theories regarding safety, authors have suggested that system safety in general is intertwined with some factors, including the redundancy of components that can perform each critical function and the existence of fail-safes⁸⁹. Elevator brakes are a classic example of a fail-safe feature. They are attached to the outside of the cabin and are held open only by the tension in the cables that the cabin is suspended on. If tension is lost in the cables, the brakes automatically clamp shut onto the rails in the elevator shaft. This means that, even if the cables break, the brakes should prevent the cabin from falling; even if the system fails in its function, it should at least be safe.

While easy cases certainly exist⁹⁰, relevant stakeholders⁹¹ urged the EC to clarify this notion, so as to prevent an expansive interpretation of the h-AIS category (see §3.1 *supra*), and feedback is expected on this very point to be found in the forthcoming guidelines on art. 6, due on February 2nd, 2026⁹².

4.3. Safety Components and Complex Systems

Finally, even if the notion of safety component were clear, it would be most complicated to understand when, indeed, a given application is the safety component of another, in all those systems that comprise more

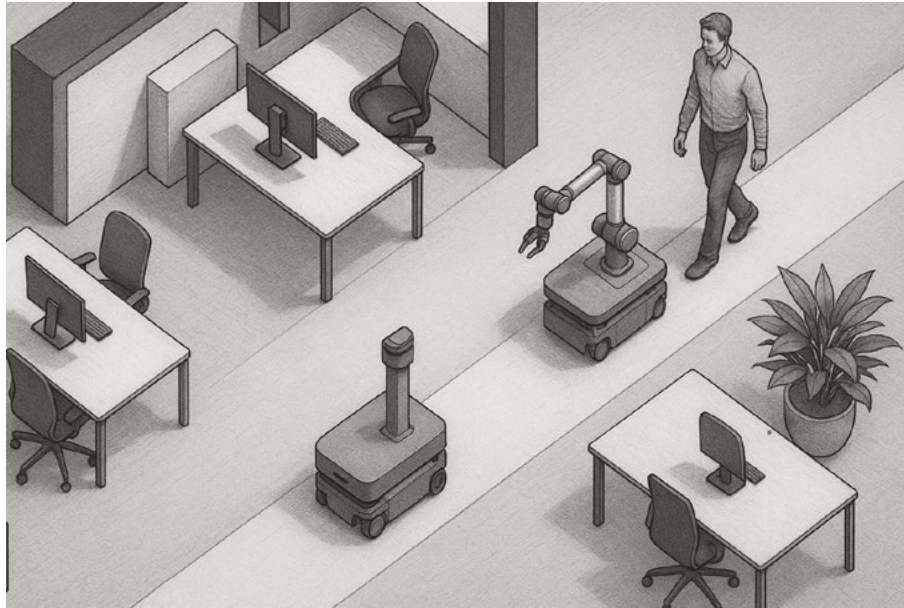
⁸⁹ D. Hendrycks, *Introduction to AI Safety, Ethics, and Society*, 2024, 188.

⁹⁰ A prime example Microsoft's project InnerEye, which uses machine learning for automatic delineation of healthy anatomy and tumors. Microsoft Research, *Project InnerEye - Democratizing Medical Imaging AI*, in [microsoft.com](https://microsoft.com/research/projects/inner-eye); For an overview on the academic production on the use of various AIS on medical diagnosis refer to S. Kaur – others, *Medical Diagnostic Systems Using Artificial Intelligence (AI) Algorithms: Principles and Perspectives*, in *IEEE*, 2020.

⁹¹ Industries regulated by the RED have manifested their discontent with the expansive interpretation provided by the EC. L. Bertuzzi, *Device Makers Using AI in Safety Features Could Have Strict AI Act Obligations, EU Commission Says*, in mlex.com.

⁹² L. Bertuzzi, *Product Makers to See AI Act Interplay with EU Safety Rules in Future Guidelines*, in mlex.com.

than one machine in themselves, or anyway see the collaboration between multiple machines. Indeed, Industry 4.0 scenarios⁹³, based in the adoption of embedded and non-embedded applications, are going to heavily rely on those forms of technological solutions, heavily relying on integration. To exemplify this concept, please consider the following case where more co-bots (collaborative robots) are allowed to roam in an office space or factory, while connected to a radar that uses an algorithm to detect and categorize potential obstacles, thence ultimately allowing the robots to avoid collision (see Figure 1).



This setup raises important questions: should the radar and processing unit alone be classified as the h-AIS? Should the robots also be included in this classification? Or should the entire integrated system—robots and radar together—be evaluated under art. 6 of the AIA? The outcome of this classification has significant implications, particularly in identifying which parties are subject to the obligations outlined in art. 8 and beyond of AIA. One conservative way of interpreting it is that only the external unit, if sold separately, is the safety component. Everything else, even if it implements safety functions, is machinery. However, the opposite interpretation is also plausible.

A separate entity, known as the system integrator, might then combine these components to meet a client's specific needs. If each robot were considered a high-risk system, the number of responsible parties would grow considerably. If, on the other hand, only the radar and processing system were deemed high-risk, responsibility would fall solely on their developers, providers, and deployers. In a third scenario, where only the integrated system as a whole is treated as high-risk, liability would rest primarily with the system integrator. This ambiguity once again highlights the serious challenges posed by the current regulatory uncertainty.

⁹³ H. Lasi – Others, *Industry 4.0*, in *Business & Information Systems Engineering*, 2014, 239 ff.

4.4. The Certification of the AIS or of its Safety Component: an Overview

Even if the system using AI was indeed a safety component, for the AIS to be deemed high risk the AIA sets an additional requirement, namely that of that very component—or the overall AIS—being subject to a TPCA under one of the listed pieces of European product safety legislation (UHL, see Annex I).

However, in order to dive into the interpretation of art. 6, para. 1, lit. b) it is necessary to clarify some essential concepts about product certification, a requirement under the so called the New Approach to Market Integration of the European Union, intended to ensure the essential safety of all products before they are distributed within the EU⁹⁴.

The process begins with identifying the applicable directive or regulation for a given product. The EU has established multiple directives and regulations that define specific requirements for different categories of goods. For example, the MD applies to industrial and consumer machinery, LVD governs electrical equipment, and the MDR ensures the safety and performance of medical devices. Other directives include those related to toys, electromagnetic compatibility, and radio equipment. A manufacturer must first determine which piece of European legislation applies to their product before proceeding further⁹⁵.

Then the manufacturer must assess the conformity route required to demonstrate compliance. Depending on the product type and its associated risk, certification may be achieved through self-assessment or require the intervention of a Notified Body (henceforth NB)⁹⁶. Products such as a LED lamp under the LVD can often be self-certified (art. 12-14 and Annex I, LVD). However, items with critical safety requirements, like a life-supporting medical ventilator, require third-party verification to ensure their safety and reliability (art. 52, para. 1, and Annex VIII MDR).

Following this, a comprehensive conformity assessment must be carried out. This process involves conducting a risk assessment to identify potential hazards, testing the product—most commonly, even if not necessari-

⁹⁴ Your Europe, *CE Marking*, in *europa.eu*.

⁹⁵ G. Ballor, *CE Marking, Business, and European Market Integration*, in *Business History Review*, 2022, 77 ff.

⁹⁶ NBs are independent organizations designated by EU member states to assess the conformity of certain products before they enter the EU market, ensuring compliance with safety, health, and environmental standards. These bodies are authorized under specific EU directives and regulations, often for high-risk products like medical devices (art. 35 MDR) and machinery (art. 30 MR). NBs play a key role in the CE marking process, which signifies that a product meets EU regulatory requirements. Their primary function is to perform conformity assessments which may involve testing, inspecting, and auditing products or manufacturing processes. The AIA has followed previous EU regulations and directives and introduced the role of NB in its Section 4. They are responsible for assessing «the conformity of high-risk AI systems in accordance with the conformity assessment procedures set out in art. 43» (art. 34, para. 1 AIA). The AIA also reproduces the regulatory standard of linking the CE mark to the certification performed by the NBs (art. 48, para. 4, AIA). See also J.-P. Galland, *The Difficulties of Regulating Markets and Risks in Europe through Notified Bodies*, in *European Journal of Risk Regulation*, 2013, 365 ff.

ly—against harmonized standards, and compiling the necessary technical documentation⁹⁷. The technical file should include a detailed product description, information about its design and manufacturing process⁹⁸, results from risk analysis, and test reports proving compliance with the relevant EU standards⁹⁹. For example, a children’s toy must undergo rigorous mechanical and chemical safety assessments to ensure it does not present a choking hazard and is devoid of toxic substances¹⁰⁰.

A TPCA will differ from the self-certification in the involvement of a NB, that will be responsible for «one or several of the following activities: calibration, testing, certification and inspection»¹⁰¹. The extent to which a notified body will be involved depends on the module of certification that is being used—as per the modules defined in Decision No 768/2008/EC. Furthermore, each piece of legislation can adopt changes to those modules, in order to better reflect the necessities of the regulated product. For instance, an EU-type examination for the RED has taken the following shape: «a conformity assessment procedure in which a notified body examines the technical design of the radio equipment and verifies and attests that the technical design of the radio equipment meets the essential requirements set out in Art. 3» (Annex III, RED). The MDR adopts a similar definition, but specifies, amongst other things, that «review the clinical evidence presented by the manufacturer in the clinical evaluation report» (Annex X, para. 3, lit. c), MDR). In sum, for the execution of a TPCA, a notified body will follow the guidelines provided in the specific legislation, that reflect the Decision No 768/2008/EC.

Once compliance has been established, the manufacturer must prepare and sign a European Declaration of Conformity¹⁰². This formal document affirms that the product meets all applicable EU directives and/or regulations, and must be retained for at least ten years¹⁰³. The next step is affixing the CE marking on the product, which signifies that it conforms to legal safety requirements¹⁰⁴. For instance, a power drill that meets the requirements of the MR must display the CE marking on its body or packaging before being placed on the market (art. 24, para. 2, MR).

After a product enters the market, manufacturers must ensure ongoing compliance. This includes post-market surveillance, addressing any safety concerns that arise, and making necessary updates to the technical documentation in case of design modifications¹⁰⁵. If a defect is identified in a product, such as a malfunction in a medical syringe pump covered under the MDR, the manufacturer is responsible for initiating a recall and noti-

⁹⁷ Your Europe, *Conformity Assessment*, in europa.eu.

⁹⁸ S. Rajput, *EU MDR and IVDR Conformity Assessment Guide*, Celegence, 2024.

⁹⁹ Your Europe, *Technical Documentation and EU Declaration of Conformity*, in europa.eu.

¹⁰⁰ European Council, *Toy Safety*, in consilium.europa.eu.

¹⁰¹ European Commission, *The “Blue Guide” on the implementation of EU products rules*, Brussels, 2016.

¹⁰² Your Europe, *Technical Documentation*, cit.

¹⁰³ *Ibid.*

¹⁰⁴ Your Europe, *Conformity Assessment*, cit.

¹⁰⁵ Your Europe, *Technical Documentation*, cit.

fying regulatory authorities (art. 10, para. 12 MDR).

Despite the comprehensive nature of the certification process, certain products are exempt from CE marking requirements. This might be the case both for products that are not covered by EU directives and may require national approvals instead, as well as custom-made equipment, designed solely for in-house use within a company, which is not made available on the open market. Furthermore, military equipment and other specialized items used exclusively for national defence purposes are often exempt from CE marking requirements.

4.5. When is a TPCA “Required”?

Art. 6, para. 1, lit. b), AIA, formulating the second condition for an AIS to be deemed high-risk, recites:

«the product [...] is required to undergo a third-party conformity assessment [...] pursuant to the Union harmonisation legislation listed in Annex I».

The choice of the verb “required”¹⁰⁶ appears sufficiently clear in restricting the scope of h-AIS to those instances where a TPCA is mandated, and not merely optional, among alternative certification techniques.

The Commission, however, hinted that it might adopt an expansive interpretation to encompass all such cases where a TPCA is merely an option for compliance, but the choice is left with the producer, who may resort to alternative solutions to certify its product, such as self-certification¹⁰⁷. Such cases are, indeed, all but uncommon¹⁰⁸.

Debates already occurred with respect to the RED and TSD respectively. As per the former, the Commission maintained that demonstrating com-

¹⁰⁶ Similar considerations may be drawn taking other official translations into account, such as the Italian—where «è soggetto a una valutazione della conformità da parte di terzi», the French—«est soumis à une évaluation de conformité par un tiers», and the German—«muss einer Konformitätsbewertung durch Dritte»—where it is clear that the TPCA is not optional for the manufacturer to choose among potential alternative certification procedures. In other instances, instead, (See for instance the Italian translation of art. 6, para. 3, co. 3), divergences in the translations and their potential meaning may be observed, that have a relevant bearing on the interpretation and possible applications of the norm. For a discussion please allow reference to A. Bertolini, *La definizione di sistemi di intelligenza artificiale*, in S. Orlando – G. Cerrina Feroni (eds.), *Il regolamento europeo sull'intelligenza artificiale: problemi e sistema*, forthcoming 2025. It shall also be noted that the same interpretation was reached by stakeholders regulated by the RED. L. Bertuzzi, *Device Makers Using AI in Safety Features Could Have Strict AI Act Obligations*, *EU Commission Says*, cit.

¹⁰⁷ The EC has elucidated its stance on the matter through the EU AI Act Compliance Checker, the novel instrument of the AI Office. Once the self-assessment has begun and you have reached the question of whether the AIS can be considered high risk pursuant to art. 6, para. 1 AIA, you must specify whether the product, or the product in which the AIS is intended to be used as a safety component, is required to undergo a TPCA under sectoral law. Well, it is precisely at this point that the EC expressly warns that this also includes products for which you can opt out of a TPCA when harmonized standards are fully applied. See ai-act-service-desk.ec.europa.eu.

¹⁰⁸ See, for instance, art. 25, para. 3, lit. a) MR; art. 20, para. 2, and 22, para. 2, TSD and art. 17, para. 3, RED.

pliance through harmonized standards, when conceived as an alternative to a TPCA (art. 17, para. 2, lit. 3), RED), would not *per se* exclude the qualification of the system as h-AIS¹⁰⁹. As per the latter, the proposed Toy Safety Regulation (TSR), expressly clarifies that «The choice by the manufacturer of the conformity-assessment procedures for such toys, where there is a possibility to opt out of third-party conformity assessment where harmonized standards have been applied, should not affect this classification as high-risk AI system», causing manufacturers to be raise concerns about the impact of the regulation¹¹⁰.

The Commission has also spoken regarding the interactions of the AIA, the MDR and the IVDR. It has issued a Q&A document¹¹¹ aimed at clarifying how those three pieces of EU law should interact. The document presented the table below, which seems to indicate that most of the products regulated by those legislations are to be considered h-AIS. The Q&A document also explicitly notes that, according to Recital 51 AIA, the fact that a product is considered a h-AIS, does not mean it will be considered high-risk for the purposes of the MDR and IVDR¹¹². For instance, for the MDR devices are classified according to their level of risk (Class I, Class IIa, Class IIb and Class III)¹¹³. Class IIa devices are considered to be medium risk, however they still need to undergo a third party conformity assessment¹¹⁴, thus being subject to h-AIS regime.

Classification	Notified Body Involved?	AIA High-Risk (Art. 6(1)) conditions fulfilled?
MDR Class I (non-sterile, non-measuring, non-reusable surgical)	✗ No	✗ No
MDR Class I (sterile, measuring, reusable surgical)	✓ Yes	✓ Yes
MDR Class IIa, IIb, III	✓ Yes	✓ Yes
MDR Annex XVI ¹⁰	✓ Yes	✓ Yes
IVDR Class A (non-sterile)	✗ No	✗ No
IVDR Class A)	✓ Yes	✓ Yes
IVDR Class B, C, D	✓ Yes	✓ Yes
In-house device according to Art. 5(5) MDR/IVDR	✗ No	✗ No

¹⁰⁹ L. Bertuzzi, *Device Makers Using AI in Safety Features Could Have Strict AI Act Obligations, EU Commission Says*, cit.

¹¹⁰ L. Bertuzzi, *Manufacturers of Toys Using AI Set to Face Strict AI Act Regime (Correct*)*, in *mlex.com*.

¹¹¹ Directorate-General for Health and Food Safety, *MDCG 2025-6 - FAQ on Interplay between the Medical Devices Regulation & In vitro Diagnostic Medical Devices Regulation and the Artificial Intelligence Act (June 2025)*, in *health.ec.europa.eu*.

¹¹² MLEX, *EU Q&A Document on AI Act, Medical-Device Rules Interplay up for Final Approval*, in *mlex.com*.

¹¹³ European Commission, *Medical device classification*, in *webgate.ec.europa.eu*.

¹¹⁴ Art. 52, para. 6, MDR determines that «Manufacturers of class IIa devices, other than custom-made or investigational devices, shall be subject to a conformity assessment as specified in Chapters I and III of Annex IX». On its turn Annex IX, Chapter I and III require the involvement of a notified body. Which, as we have seen, triggers the application of art. 6, para. 1.

If required were to mean an inescapable legal obligation—as the most logical interpretation of the AIA would suggest—the result would be that only a limited amount of products regulated under UHL would be considered high-risk, as those directives and regulations provide different avenues through which the conformity of a product can be assessed, some of which, do not involve the use of notified body, thence a third-party conformity assessment. Instead, if we were to consider the term required to include instances where resorting to NB is among the possibilities, then every product under a UHL would have to be considered high-risk. As a matter of example, let's look into the MR and how it has employed third-party conformity assessments.

Art. 25 MR establishes that products that fall under its scope will be regulated according to procedures (called modules) designated in its annexes:

- «2. Where the category of machinery or related product is listed in Annex I, Part A, the manufacturer or the natural or legal person referred to in art. 18 shall apply one of the following procedures:
 - (a) EU type-examination (module B) set out in Annex VII, followed by conformity to type based on internal production control (module C) set out in Annex VIII;
 - (b) conformity based on full quality assurance (module H) set out in Annex IX;
 - (c) conformity based on unit verification (module G) set out in Annex X.
- 3. Where the category of machinery or related product is listed in Annex I, Part B, the manufacturer or the natural or legal person referred to in art. 18 shall apply one of the following procedures:
 - (a) internal production control (module A) set out in Annex VI;
 - (b) EU type-examination (module B) set out in Annex VII, followed by conformity to type based on internal production control (module C) set out in Annex VIII;
 - (c) conformity based on full quality assurance (module H) set out in Annex IX;
 - (d) conformity based on unit verification (module G) set out in Annex X.»

Clearly, art. 25 MR allows an option to the manufacture to choose between the modules, while still being able to demonstrate compliance with the regulation. Among the different modules recalled, only B, H and G unequivocally require the intervention of a notified body, whereas A establishes a procedure that is entirely based on the assessment of the manufacturers themselves¹¹⁵. Taking that into account, we can observe how all the options enumerated under para. 2 require the intervention of a NB, whilst para. 3 ensures the manufacture the possibility to resort to self-certification, while still maintaining the choice to opt for a TPCA. Provisions like that appear throughout all the UHL¹¹⁶. If we were to argue that the TPCA

¹¹⁵ Art. 25 MD determines that, regarding Module A, «the manufacturer fulfils the obligations laid down in points 2, 3 and 4, and ensures and declares on its sole responsibility that the machinery or related product concerned satisfies the applicable requirements of this Regulation».

¹¹⁶ Art. 12 MD; art 25, MR; art. 19 TSD; art. 20, para. 2, and 22, para. 2, Watercraft

is required only when the manufacturer has no option but to conduct such a procedure, this interpretation might be considered too restrictive and that it would leave effectively high-risk products outside the application the AIA. However, such criticism to the literal interpretation of art. 6, para. 1, would appear disproportionate, in as much as those products are already heavily regulated under UHL.

4.6. Case Studies: on the Marginal Discrepancies

Building on the analysis carried out above (see §§ 4.1-4.5), we have conceived a series of three examples of AI-based Systems that are already existent or currently under research. We have construed those examples taking into account the respective technologies, and carried out individual interviews with experts to allow us to determine the correct functioning of those applications, possibly identify their respective application and help us identify what may be deemed a safety component in the specific examples at hand, and determine whether it would be subject to the UHL indicated in Annex I.

Those examples include, a security patrol robot manufactured according to hENs, the humanoid robot with comparable characteristics to that of Boston Dynamics's Atlas¹¹⁷, and a drone docking station technology in personal watercrafts.

Firstly, a security patrol robot is deemed to be a (semi-)autonomous mobile robot that repeatedly visits or patrols specified areas, equipped with sensors, cameras, and decision-making software, to detect anomalies—such as motion, intrusions, or environmental hazards—and either respond directly or transmit alerts for further action¹¹⁸.

These robots are typically equipped with a range of advanced sensors, including but not limited to RGB¹¹⁹ and infrared cameras, microphones, ultrasonic range finders, and environmental sensors (e.g. for temperature, smoke, motion, and light), which enable them to perceive their surroundings, identify suspicious behaviour, and detect intrusions or hazardous conditions.

From a technological standpoint, such robots are capable of autonomous navigation, relying on GPS and RTK¹²⁰ systems in outdoor environments

Directive; art. 13, para. 1, Directive 2014/34/EU; art. 17, para. 2, lit. a) RED; art. 14, para. 2, lit. a) and b) Pressure Equipment Directive; art. 19, para. 1 PPE Regulation.

¹¹⁷ None of the considerations here drawn are to be intended as a binding legal analysis of the specific robot by Boston Dynamics which is, instead, intended as a plausible and realistic example of a technology that could fall under the scope of application of the AIA.

¹¹⁸ N. Basilico, *Recent Trends in Robotic Patrolling*, in *Current Robotics Reports*, 2022, 65–76.

¹¹⁹ RGB stands for Red, Green, Blue—the three primary colors of light used in digital imaging and display technologies. In the context of RGB cameras, this term refers to standard digital cameras that capture visual information by detecting the intensity of red, green, and blue light in a scene.

¹²⁰ RTK stands for Real-Time Kinematic positioning. It is a satellite navigation technique used to enhance the precision of position data derived from global navigation satellite

and on Simultaneous Localization and Mapping (SLAM) algorithms or track-based navigation for indoor deployment. The decision-making component of these systems is often based on AI techniques, such as convolutional neural networks for facial recognition and object detection, and fuzzy logic or machine learning algorithms for route optimization, obstacle avoidance, and behaviour analysis¹²¹.

In order to determine the classification of the AIS as high risk or not, as outlined in art. 6, para. 1, AIA, it was necessary to legally isolate the safety component.

In accordance with the insights provided by subject matter experts, it is imperative to conceptualize the safety system in its entirety, rather than focusing on its constituent components.

At this juncture, it was necessary to ascertain whether any of the UHLs enumerated in Annex I pertained to this particular robot and whether it was subject to conformity assessment by a notified body (TPCA).

It is our opinion, based on the available evidence, that the MD apply to this robot by virtue of para. 19, Annex IV MD (“protective devices designed to detect the presence of persons”).

The aforementioned directive enumerates a series of cases in which a TPCA is requested but only art. 12, para. 4, MD relates to h-AIS in the sense of our perspective (option A) – in other words, the term “required” indicates something necessary and indispensable. Therefore, only AIS that are mandatorily subject to TPCA can be considered high risk. Consequently, a security patrol robot that has not been manufactured in accordance with the hENs and that is subject to EC type-examination or full quality assurance procedure can be classified as high risk (see art. 12, para. 4, MD).

Nevertheless, an issue emerges when the security patrol robot is produced in compliance with the hENs and the latter encompass all pertinent essential health and safety requirements, as stipulated in art. 12 MD. In this particular instance, manufacturers are not obliged to follow a TPCA and, in accordance with our interpretation of the term “required,” the robot is not deemed to pose a significant risk.

This perspective is not shared by those who contend that the term “required” signifies a TPCA that is not essential but rather one of the potential methods of conformity assessment that can be implemented (option B).

Secondly, we analysed a humanoid robot with comparable characteristics to Atlas. As that of Boston Dynamics it is a bipedal humanoid robot primarily designed as a research platform for studying agile mobility, human-like movement, and robot-environment interaction in unstructured environments. The robot exemplifies advanced capabilities in dynamic locomotion, real-time perception, and adaptive control. Its design aims to approximate the morphology and movement of the human body, there-

systems (GNSS), such as GPS.

¹²¹ For an overview on security patrolling robot from a technical point of view please see H. Liu – L. Yu – W. Meng – Q. Zhu – Y. Wang, *Patrol planning for a security robot using a spatio-temporal model*, in *Robotics and Autonomous Systems*, 92, 2017, 16–26; J. Choi – D. Kim, *Design and Implementation of an Indoor Patrol Robot System*, in *International Journal of Advanced Robotic Systems*, 17, 2020, 1–11.

by enabling it to perform complex tasks in unstructured environments. The robot is approximately 1.5 meters tall and weighs around 80 to 90 kilograms, dimensions that make it suitable for navigating human-centric spaces such as buildings, staircases, or rough terrain.

Besides the mechanical architecture, which includes 28 degrees of freedom (DoF), the robot integrates an advanced sensor suite to perceive and interpret its environment. This includes stereo vision cameras, LiDAR sensors, and an inertial measurement unit (IMU). These sensors feed into a real-time perception system that constructs a three-dimensional model of the surrounding environment, which is essential for obstacle negotiation, path planning, and terrain adaptation. Visual-inertial odometry enables Atlas to estimate its position and orientation even in the absence of GPS or external tracking systems. Another defining feature is its control architecture, which combines classical control techniques with modern machine learning and model predictive control (MPC). The robot uses, indeed, AI-based algorithms for real-time perception, which are essential for interpreting complex environments¹²².

The classification of robots as high risk is contingent upon the delineation of the safety component's perimeter. In light of the expert advice received, the entire system was identified.

At this juncture, the pertinent UHL was ascertained. Consequently, the MD were deemed applicable. In this instance as well, reference should be made to art. 12, para. 4, of the MD by virtue of para. 17 of Annex IV MD ([d]evices for the lifting of persons or of persons and goods involving a hazard of falling from a vertical height of more than three meters).

Given the robot's status as an American robot—as that of Boston Dynamics—for which European production standards are not obligatory nether suggested, its noncompliance with these standards would, in our estimation (option A), readily categorize it as a high-risk Ais if undergoing a TPCA.

Conversely, if even a subset of the European standards were to be observed, the conclusion would be reversed. In accordance with the provisions stipulated in the relevant legislation, Atlas would not be classified as a high-risk entity unless the adoption of (option B) were to be implemented.

Lastly, we took into account a drone docking station technology in personal watercrafts. A drone docking station technology in personal watercrafts (PWCs) refers to an integrated system that allows a drone (usually an aerial drone) to autonomously land, recharge, and possibly transmit data directly from the watercraft. This innovation combines unmanned aerial vehicle (UAV) capabilities with marine vehicles, enhancing their functionality, especially for recreational, rescue, and surveillance applications. It includes a stabilized landing platform designed to compensate for the motion of the watercraft, ensuring safe and accurate autonomous landings. The station integrates a wireless or contact-based charging mechanism powered by the PWC's onboard electrical system, allowing the drone to recharge

¹²² Since the robot's characteristics are comparable to those of Atlas, for further technical details, please refer to Boston Dynamics, [Atlas® and beyond: the world's most dynamic robots](https://www.bostondynamics.com/atlas/), in [bostondynamics.com/atlas/](https://www.bostondynamics.com/atlas/).

without manual intervention. A data transfer module, typically utilizing Wi-Fi, Bluetooth, or USB-C, enables seamless uploading of aerial footage or sensor data to onboard storage or remote servers. The system relies on GPS and AI-assisted navigation algorithms to support autonomous take-off, tracking, and return-to-dock functions, even while the watercraft is in motion. Additionally, the docking station is housed within a marine-grade protective enclosure to resist water, salt, and vibrations, ensuring durability and operational reliability in harsh aquatic environments¹²³.

The safety component—even here—can be deemed to be the whole system while the applicable UHL is the Directive 2013/53 on recreational craft and personal watercraft (RCPW). Being the example a personal watercraft, the referral is to art. 20, para. 2, RCPW which provides various type of conformity assessment including TPCA.

According to our reasoning, this technology cannot be qualified as h-AISs given that it does not stipulate TPCA as the sole option for conformity assessment (option A).

Nevertheless, the alternative interpretation (option B) posits that the drone self-docking station in personal watercraft could be classified as h-AIS. This is predicated on the premise that the legislation in effect allows for the subjecting of the product to TPCA.

As illustrated by the provided examples, the differentiation between h-AISs and AISs is predicated on marginal discrepancies. In the initial two scenarios, the distinction is determined by adherence to heN, which are not mandatory. In the latter scenario, the distinction primarily lies on the quality of personal or recreational watercraft.

¹²³ For an overview on technical features on drone self-docking station in personal watercrafts see C. G. Grlj – N. Krznar – M. Pranjić, *A Decade of UAV Docking Stations: A Brief Overview of Mobile and Fixed Landing Platforms*, in *Drones*, 2022.

AI System example	Interpretation of «required» - option <u>A</u>	Interpretation of «required» - option <u>B</u>
Security patrol robot* <i>Para. 19, Annex IV, Directive 2006/42/EC¹²⁴ as recalled in art. 12, para. 4¹²⁵ same Directive</i>	<i>NOT high-risk</i>	<i>high-risk</i>
Humanoid robot* <i>Para. 17, Annex IV, Directive 2006/42/EC¹²⁶</i>	<i>NOT high-risk</i>	<i>high-risk</i>
<u>Drone docking station technology in personal watercrafts</u> (<i>eg. GODO A170 Dock</i>) <i>Art. 20, para. 2, Directive 2013/53/EU¹²⁷</i>	<i>NOT high-risk</i>	<i>high-risk</i>

Legenda:

Option A: «required» is intended in a strict way. It stands for “TPCA is compulsory”.

Option B: «required» is intended in a broad way. It stands for “TPCA is possible”.

Method: Examples are extracted from each recalled directive.

*The product is equipped with AI capabilities. In the context of h-AISs qualification, components are regarded as an integrated whole.

¹²⁴ Para. 19, Annex IV Directive 2006/42/EC of the European Parliament and of the Council of 17 May 2006 on machinery, and amending Directive 95/16/EC, OJ L 157, 9.6.2006: «Protective devices designed to detect the presence of persons».

¹²⁵ art. 12, para. 4, *Ibid.* «Where the machinery is referred to in Annex IV and has not been manufactured in accordance with the harmonized standards referred to in art. 7, para. 2, or only partly in accordance with such standards, or if the harmonised standards do not cover all the relevant essential health and safety requirements or if no harmonised standards exist for the machinery in question, the manufacturer or his authorised representative shall apply one of the following procedures: (a) the EC type-examination procedure provided for in Annex IX, plus the internal checks on the manufacture of machinery provided for in Annex VIII, point 3; (b) the full quality assurance procedure provided for in Annex X».

¹²⁶ Para. 17, Annex IV Directive 2006/42/EC cit.: «Devices for the lifting of persons or of persons and goods involving a hazard of falling from a vertical height of more than three meters».

¹²⁷ European Parliament and the Council, Directive 2013/53/EU of the European Parliament and of the Council of 20 November 2013 on recreational craft and personal watercraft and repealing Directive 94/25/EC, OJ L 354, 28.12.2013, 90-131, art. 20, para. 2.

5. A Critical Perspective

In conclusion, the interpretation of art. 6, para. 1, lit. b) of the AIA plays a pivotal role in determining the scope of h-AIS under the new regulatory regime. The crux of the issue lies in the understanding of the term “required” in relation to TPCAs under UHL. A literal reading confines the AIA’s applicability to those instances where a TPCA is legally mandatory. This interpretation maintains regulatory clarity and respects the division of responsibilities between the AIA and sector-specific UHL regimes, which have long governed conformity assessment procedures.

However, the Commission’s more expansive approach raises significant concerns. Such an interpretation may lead to the overclassification of AI products as high-risk, even where manufacturers legally opt for less stringent conformity pathways. Examples drawn from the RED, TSD, MDR, IVDR, and the MR reveal how prevalent such optional structures are, and how the Commission’s broad reading could inadvertently result in an overly burdensome regulatory landscape for AI developers and manufacturers. Moreover, expanding the definition of “required” to encompass optional conformity routes risks undermining legal certainty and predictability. As already seen in the case of stakeholders regulated by the RED, manufacturers are left uncertain as to whether choosing a less burdensome, yet lawful, certification method might still expose them to additional regulatory obligations under the AIA.

6. The Unsolved Riddle of Art. 6 AIA

In the myth of Oedipus, the riddle posed by the Sphinx was ultimately resolved by identifying a coherent metaphor for the stages of human life. By contrast, the riddle embedded in art. 6 of AIA resists such elegant resolution. Rather than culminating in a unifying interpretation, the regulation’s provisions surrounding h-AIS remain fragmented and opaque. Like the multiple heads of a hydra or the distorted echoes in a labyrinth, each interpretative effort to define what constitutes a h-AIS under art. 6 gives rise to new ambiguities and regulatory contradictions. What emerges is not a clarified standard, but a spectrum of possible readings that fail to converge on a single, stable answer.

The core of this ambiguity lies in the complex and sometimes contradictory architecture of art. 6 itself, which is divided into distinct yet overlapping prongs: the automatic classification via Annex III (para. 2), the exemptions framed in vague and elastic language (para. 3), the peculiar mechanism of provider self-declaration (para. 4), and the reference to harmonised product legislation (para. 1). Far from offering a harmonised framework, these provisions create a regulatory palimpsest—a layering of legal instruments, exceptions, and interpretative leeway that renders the classification of AIS an exercise in probabilistic reasoning rather than legal certainty. One cannot, *ex ante*, determine with confidence whether a given AIS will be deemed high-risk without navigating a thicket of technical, functional, and semantic variables, often contingent on discretionary as-

assessments by both private and public actors.

The legislative intent to adopt a technology-neutral approach further compounds this uncertainty. The AIA aims to regulate the use of AI regardless of its technical design, foregrounding the notion of risk to fundamental rights, health, and safety as the primary classificatory axis. Yet, by grounding the high-risk classification in Annex III—a list of sector-specific use cases defined not by demonstrable metrics of harm but by broad socio-political considerations—the regulation departs from genuine technology neutrality. It institutes a *de facto* taxonomy of “acceptable” versus “concerning” AI applications, without articulating a transparent risk assessment methodology or a coherent theory of harm. The very notion of “risk” becomes a rhetorical device rather than a measurable criterion. As such, systems are not categorised on the basis of demonstrated or anticipated risk levels, but through a normative framework lacking analytical precision or empirical grounding.

This conceptual fragility is further evidenced in the self-declaration mechanism of art. 6, para. 4, which ostensibly empowers providers to argue that their systems are not high-risk despite falling within Annex III. On the surface, this provision introduces a welcome degree of flexibility and contextual nuance. In practice, however, it exacerbates legal uncertainty. It does so by failing to establish clear criteria, procedures, or authoritative review mechanisms that could validate such claims *ex ante*. Providers are left to navigate a compliance landscape without a reliable compass, while regulators retain discretion to override these determinations retroactively. This asymmetry not only undermines legal certainty but also risks eroding trust in the regulatory framework as a whole.

Equally problematic is the vague terminology embedded in art. 6, para. 3, which outlines exemptions based on the nature and influence of an AIS's function. Phrases such as «narrow procedural task», «merely prepares input», or «not materially influencing» are inherently contestable and do not benefit from operational definitions. Their application requires context-sensitive judgments that are unlikely to yield uniform outcomes across Member States. This is particularly true in domains where the boundary between assistance and decision-making is blurred by the persuasive authority of algorithmic output—such as in healthcare diagnostics or judicial support tools. The regulation's reliance on such open-textured terms invites divergent national interpretations and fragmented enforcement practices, running counter to the harmonisation objectives of an EU regulation.

Finally, para. 1, which is intended to be the most technical provision, directly linking qualification as an h-AIS to the intrinsic characteristics of the system rather than embedded policy evaluations, still leaves considerable room for interpretation. Indeed, it refers to a potentially ambiguous concept - that of a safety component - and adopts wording that could either lead to an extremely broad application of the h-AIS category or, on the contrary, to a practical abrogation (at least as per para. 1).

The ambiguity of art. 6 is not incidental; it is the systemic by-product of an attempt to graft risk-based regulation onto a technology-neutral legislative foundation. The resulting synthesis, while politically palatable,

is doctrinally unstable and its implementation remains uncertain. A true risk-based framework would necessitate the identification of risk vectors, quantification of potential harm, and an evidence-based methodology for ranking and mitigating those risks. Conversely, a genuinely technology-neutral regulation would require abstract, function-agnostic rules of general applicability, applicable regardless of context. The AIA accomplishes neither. Its hybrid architecture results in a definitional impasse: a system that aspires to universality but defaults to categorical exceptions; that aims to mitigate risk but lacks evaluative tools to determine its thresholds; that seeks harmonisation but breeds interpretative fragmentation. Thus, unlike Oedipus, the reader of art. 6 cannot solve the riddle by grasping its inner logic—for it has none that is both coherent and complete. The Brussels Sphinx offers no closure but rather ultimately functions as a standing executive oracle for the provisions of the AIA, which carries the potential for bureaucratic proliferation and ambiguity in its implementation. Indeed, it leaves us suspended in a state of radical uncertainty, where developers, regulators, and stakeholders must perpetually renegotiate the boundaries of compliance. This condition is not merely an epistemic inconvenience; it is a structural vulnerability. It jeopardises the rule of law values that EU legislation is meant to uphold – legal certainty, foreseeability, and uniform application – while simultaneously disincentivising innovation and chilling market entry. Until these ambiguities are addressed through interpretative guidance, delegated acts, or eventual judicial review, art. 6 will remain less a cornerstone of AI regulation than its most impenetrable enigma.

Abstract

The article examines the conceptual and normative “riddle” posed by art. 6 of the EU Artificial Intelligence Act (AIA) in defining “high-risk” AI systems (h-AISs). It argues that the combination of a horizontal, technology-neutral framework with a risk-based classification generates significant interpretative uncertainty and undermines legal certainty. After situating the AIA within the broader EU product-safety regime and the New Legislative Framework, the contribution meticulously examines in detail the critical issues arising under Art. 6 AIA. These range from para. 2 recalling the Annex III list of high-risk AI systems, which does not rest on an objective assessment of risk, to the exceptions in paras. 3 and 4, and the cross-reference to Union harmonisation legislation in Annex I. Particular attention is paid to contested notions such as “safety component” and “third-party conformity assessment required”, illustrated through case studies (e.g. security mobile robots, humanoid robots, drone docking stations). The article concludes that this unstable definitional architecture undermines consistent application, equal treatment across sectors, and effective incentives for innovation.

Keywords

Artificial intelligence - AI Act - risk-based approach - high-risk AI systems
- interpretative uncertainties

Rule of Law in the Age of AI*

Federica Paolucci, Oreste Pollicino, Charlotte Waldron

Table of Contents

1. Introduction. – 2. Rule of Law: Theoretical Basis. – 2.1 The EU's Conception of Rule of Law: A Means, Not a Goal. – 3. Rule of Law and AI: The Challenge of Autonomy. – 4. Understanding Autonomy: The Technical Perspective. – 5. The AI Act: A Constitutional Response? – 6. Testing the Rule of Law in Context: Drawing From the Venice Commission. – 6.1 Legality. – 6.2 Legal Certainty. – 6.3 Abuse of Power. – 6.4 Equality and Non-Discrimination. – 6.5 Access to Justice. – 7. *De Iure Condendo*.

1. Introduction

«Any society in which no provision is made for guaranteeing rights or for the separation of powers has no Constitution», states art. 16 of the Declaration of Human and Civic Rights of 26 August 1789.¹

The force of this provision lies in its insistence that the law's purpose is not abstract moral vindication, but the concrete disciplining of power.² The rule of law, in this European understanding,³ does not describe a final state of affairs but a structural condition: it ensures that authority is channelled through general and foreseeable norms, and that it remains subject to institutional checks capable of providing redress.⁴ Its function is instru-

* L'articolo è stato sottoposto, in conformità al regolamento della Rivista, a referaggio "a doppio cieco".

This article reflects the views of all the authors. Authorship for Sections 2, 2.1, 3 and 4 lies primarily with Oreste Pollicino, while Sections 5, 6, 6.1, 6.2, 6.3, 6.4 and 6.5 were written primarily by Federica Paolucci. The introduction and conclusion are the result of collective drafting.

¹ [Declaration of Human and Civic Rights](#), 26 August 1789.

² *Ex multis*, C. de Montesquieu, *The Spirit of Laws*, Book XI, in *Wikisource*.

³ This paper adopts a primarily European perspective on the rule of law. Comparative dimensions are acknowledged, and reference is made to A. Sajó, *The Rule of Law*, in R. Schütze-R. Masterman (eds.), *The Cambridge Companion to Comparative Constitutional Law*, Oxford, 2019, 258 ff. Sajó underlines the plurality of national traditions but also the specifically European understanding of the rule of law as a structural condition rather than a final state of affairs. Methodologically, this paper adopts an analysis that is both doctrinal and institutional. It reconstructs the Union's rule-of-law standard from primary law and the Court of Justice's jurisprudence and treats conditionality mechanisms, particularly the budgetary conditionality regulation and the Recovery and Resilience Facility, as examples of how constitutional standards are rendered administrable through verifiable safeguards.

⁴ On this point, see Marta Cartabia who highlights that, in the European constitutional

mental. It is not an end in itself,⁵ but the indispensable means by which fundamental rights and other relevant values are secured and granted. Alongside the more formalistic historical visions of the rule of law, focused on legality in its narrow sense, namely that public action be authorised by law, scholarship and institutional organisations have progressively emphasised its substantive dimension.⁶ The rule of law is not only about constraining authority through predictable and general norms but also about securing rights and access to redress.⁷ This double orientation, law as a restraint on power, and the rule of law as a means to protect rights, is precisely what explains why the debate about Artificial Intelligence (hereinafter, AI) cannot be cast in purely technological or efficiency terms. The rule of law has already faced severe challenges in past constitutional crises in Europe, where pressures to bypass checks and guarantees in the name of emergencies or expediency weakened its foundations.⁸ Those experiences show that crises of legality and rights protection are not new phenomena; the advent of AI risks reproducing them in new forms, or even exacerbating them, by introducing technological opacity and a potential displacement of responsibility.⁹ When decisions are increasingly mediated, or even taken, by automated systems, the question that arises is whether the guarantees associated with legality, separation of powers, and effective judicial protection remain in-

tradition, the rule of law is not conceived as a static outcome but as a structural requirement that channels public authority through general and foreseeable norms, while ensuring its submission to institutional checks capable of redress. See M. Cartabia, *The Rule of Law and the Role of Courts*, in *Italian Journal of Public Law*, 10, 2018, 1.

⁵ «The chief merit of the teleological approach to the rule of law is that it puts this overarching purpose into perspective. The rule of law is not a goal in and of itself that should be fetishised, but a means that can help achieve other related values, such as the protection of human rights and public accountability, and thus democracy», N. A. Smuha (ed.), *The Rule of Law*, in *Algorithmic Rule By Law: How Algorithmic Regulation in the Public Sector Erodes the Rule of Law*, Cambridge, 2024, 95 ff.

⁶ On formalistic accounts see A. Sajó, *The Rule of Law*, cit., 1, noting the risks of reducing the rule of law to mere legality. On the crises that have already tested the resilience of the rule of law in the EU see F. Casolari, *Il rispetto della rule of law nell'ordinamento giuridico dell'Unione europea: un dramma in due atti: The respect of the rule of law within the legal order of the European Union: A drama in two acts*, in *DPCE Online*, 28, 2016, 135 ff. On how technological opacity and algorithmic governance risk reproducing or aggravating these crises see N.A. Smuha (ed.), *From Rule of Law to Algorithmic Rule by Law*, in *Algorithmic Rule By Law: How Algorithmic Regulation in the Public Sector Erodes the Rule of Law*, Cambridge, 2024, 151 ff.

⁷ This aspect is particularly relevant when considering the role of the judiciary, as emphasised by leading scholarship: M. Cartabia, *The Rule of Law and the Role of Courts*, cit.; K. Lenaerts, *The Rule of Law and the Coherence of the Judicial System of the European Union*, in *Common Market Law Review*, 44(6), 2007, 1625 ff.; D. Lustig-J. H. H. Weiler, *Judicial review in the contemporary world-Retrospective and prospective*, in *International Journal of Constitutional Law*, 16(2), 2018, 315 ff.

⁸ J. H. H. Weiler, *Europe In Crisis—On «Political Messianism», «Legitimacy» and the «Rule Of Law»*, in *Singapore Journal of Legal Studies*, 2012, 248 ff.

⁹ On AI autonomy, allow the referral to O. Pollicino-F. Paolucci, *Regulating AI Autonomy: A Constitutional Framework for the Digital Era*, in M. Durante - U. Pagallo (eds.), *The De Gruyter Handbook on Law and Digital Technologies*, Berlin, 2025, 353 ff.

tact.¹⁰ The issue is not only that AI may produce errors or biases, but that it reconfigures where power is exercised and how it is justified.¹¹ Decisions that once bore the marks of human judgement, subject to established administrative or judicial procedures, now emerge from systems whose rationality is statistical, whose operations are opaque, and whose governance depends on technical as well as legal choices.¹² In this sense, AI is not external to the constitutional order but internal to it: it both tests and redefines the conditions under which the rule of law is realised.¹³

To analyse this tension, it is necessary first to clarify what is meant by the rule of law in the European context. Although scholarly debate distinguishes “thin” or formal accounts, legality, prospectivity, clarity, and stability, from “thick” or substantive accounts that fold rights and equality into the concept itself, both converge on the idea that the rule of law is a discipline of power. Raz’s conception of the rule of law as a “virtue” of legal systems, ensuring that law can guide conduct, highlights its formal dimension.¹⁴ Waldron’s emphasis on procedure and contestation underlines its practical and participatory nature.¹⁵ On a different level, the Venice Commission’s Rule-of-Law Report synthesises these traditions into five operational pillars: legality, legal certainty, the prohibition of arbitrariness, equality before the law, and access to justice.¹⁶ In the EU’s legal order, these have been rendered justiciable through art. 19(1) TEU and art. 47 CFR, which the Court of Justice has read together with art. 2 TEU to produce enforceable institutional minima.¹⁷

¹⁰ These guarantees correspond to the essential components of the rule of law as identified by the Venice Commission: European Commission for Democracy through Law (Venice Commission), *Rule of Law Checklist*, CDL-AD(2016)007, adopted 11–12 March 2016, paras 12–15.

¹¹ K. Crawford, *The atlas of AI: Power, politics, and the planetary costs of artificial intelligence*, Yale, 2021; G. De Gregorio, *The Normative Power of Artificial Intelligence*, in *Indiana Law Journal*, 30(2), 2023, 55 ff.; O. Caramaschi, *Il costituzionalismo al cospetto dell’intelligenza artificiale: nuove sfide, quali soluzioni*, in *Rivista italiana di informatica e diritto*, 7(1), 2025, 23 ff.

¹² This dynamic has been vividly illustrated in the European debate by the *SCHUEA* case, concerning algorithmic credit scoring, and by controversies surrounding the deployment of biometric recognition technologies, both of which highlight how opaque, statistical rationalities challenge the guarantees of transparency, accountability and effective judicial redress. See ECJ, Case C-634/21, *OQ v. Land Hesse*, 7 December 2023; ECJ, Case C-817/19, *Ligue des droits humains ASBL v Conseil des ministres*, 21 June 2022. From a doctrinal point of view, see P. De Hert-G. Bouchagiar, *European Biometric Surveillance, Concrete Rules and Uniform Enforcement: Beyond Regulatory Abstraction and Local Enforcement*, in *Cambridge Handbook on Facial Recognition in the Modern State*, Oxford, 2024, 139 ff.; E. Falletti-C. Gallese, *Credit scoring judicial review between the Court of Justice of the European Union and comparative case law*, in *Rivista di diritto dei media*, 3, 2024, 92 ff.

¹³ P. Burgess, *AI and the Rule of Law: The Necessary Evolution of a Concept*, in P. Burgess (eds.), *AI and the Rule of Law: The Necessary Evolution of a Concept*, Oxford, 2024.

¹⁴ J. Raz, *The Rule of Law and its Virtue*, cit.

¹⁵ J. Waldron, *The Rule of Law*, in E. N. Zalta - U. Nodelman (eds.), *The Stanford Encyclopedia of Philosophy*, Stanford, 2023.

¹⁶ See the “Report on the Rule of law”, adopted in Venice, March 25–26, 2011 - CDL-AD (2011)003rev., and the “Rule of Law Checklist”, adopted in Venice, March 11–12, 2016 - CDL-AD(2016)007.

¹⁷ ECJ, Case C-64/16, *Associação Sindical dos Juizes Portugueses (ASJP) v Tribunal de Contas*,

What matters for present purposes is that the rule of law operates as a bridge between legality and rights.¹⁸ Its formal dimension disciplines how power is exercised: laws must be clear, general, and foreseeable; decisions must be taken by authorised bodies on the basis of lawful procedures; records must be kept and reasons given. Its substantive dimension ensures that legality does not become an empty shell: decisions must respect equality and non-discrimination, remain proportionate, and allow for effective remedies before an independent tribunal. This combination of form and substance makes the rule of law the principal means by which fundamental rights are guaranteed in practice.

Artificial intelligence activates both dimensions simultaneously. Exogenously, when automated systems are used in domains such as welfare, education, migration or credit, the legality of their deployment depends on whether the decision-making process respects the principles just outlined: whether it is supported by a lawful basis and purpose, whether transparency obligations are discharged, whether human oversight is real, and whether remedies are available in time to matter.¹⁹ Endogenously, AI also raises questions about the legality of design and governance within the systems themselves: how data are sourced and managed, how objectives and thresholds are defined, how retraining and change control occur, and how responsibility is distributed along the provider–deployer chain.²⁰ These are not technical details but normative choices. They define who is accountable, what counts as an error, and which values are privileged in trade-offs between efficiency and fairness. To ignore them would be to allow a displacement of legal judgment into opaque technical parameters.²¹ The European Union has progressively operationalised this functional understanding of the rule of law through its own reporting practice. The 2025 Rule of Law Report and its updated methodology confirm that le-

27 February 2018, §31-52, where the Court held that art. 19(1) TEU, read in conjunction with art. 47 CFR and art. 2 TEU, requires Member States to ensure effective judicial protection and judicial independence as part of the rule of law.

¹⁸ «In some ways, it is once again the principles of the rule of law that are trying to rise from the ruins of the nation state and regain strength on a global scale. It is still the age-old dream of imposing rules on power. But dreaming is not enough: as has always been the case in the difficult history of the struggle for rights, dreams must be accompanied by political struggle», R. Bin, *Lo Stato di diritto*, Bologna, 2017, 120 (aut. trans.).

¹⁹ See District Court of The Hague, 5 February 2020, *NJCM c.s./Netherlands (SyRI)*, declaring unlawful the use of an automated risk-scoring system in welfare on grounds of lack of transparency and disproportionate interference with the right to privacy under art. 8 ECHR. Scholars argue on the human rights implication of this system of digital welfare, as in A. Rachovitsa-N. Johann, *The Human Rights Implications of the Use of AI in the Digital Welfare State: Lessons Learned from the Dutch SyRI Case*, in *Human Rights Law Review*, 22(2), 2022, 10 ff.; S. Ranchordás, *The invisible citizen in the digital state: Administrative law meets digital constitutionalism*, in J. De Poorter et al. (eds.), *European yearbook of constitutional law*, ebook, 2024; M. van Bekkum-F. Z. Borgesius, *Digital welfare fraud detection and the Dutch SyRI judgment*, in *European Journal of Social Security*, 23(4), 2021, 323 ff.

²⁰ On the legality of AI system design and governance, see Regulation (EU) 2024/1689 (hereinafter, AI Act), arts. 9-15, 26 (risk management, data governance, human oversight and accountability across the provider–deployer chain).

²¹ On the necessity to redefine and clarify the boundaries of the AI legal responsibility, see G. Finocchiaro, *The regulation of artificial intelligence*, in *AI & Society*, 39, 2024, 1961 ff.

gality, legal certainty, prevention of abuse of power, equality and access to justice are not rhetorical values but working benchmarks²² which are applied by the EU in specific areas: justice, anti-corruption framework, media freedom and media pluralism.²³ By translating art. 2 TEU into annual monitoring indicators, the Commission has effectively aligned itself with the Venice Commission, treating the rule of law as a set of verifiable conditions for rights protection.²⁴ This reinforces the claim that in the Union's constitutional grammar, the rule of law is a *means*, not an end.

Thus, the purpose of this paper is to assess whether the European Union's emerging framework for AI meets its own rule-of-law standard when both dimensions are taken into account. The central claim is twofold. Doctrinally, this article argues that the Union's conception of the rule of law is best understood as a functional principle: it is the means by which fundamental rights are made effective and public power is disciplined. Read in that key, the challenge of artificial intelligence is not merely whether new rules exist, but whether those rules deliver *remedy-capable* guarantees when decision-making becomes automated and partly opaque. Methodologically, the analysis anchors this assessment in the Venice Commission's Rule of Law Checklist, treating its benchmarks, legality, legal certainty, prevention of abuse of power, equality and non-discrimination, and access to justice, not as abstract values but as operational conditions that must be satisfied both in the use of AI and in its design and governance.

The paper proceeds by building this framework and then testing it against the Union's regulatory response. While Section II reconstructs the theoretical basis of the rule of law, Section III maps how autonomy in AI, understood as learning, adaptiveness and partial independence from human reasoning, reconfigures the *loci* of power and therefore the terms on which the rule of law must operate. Then, Section IV offers a concise technical account of autonomy to ground the legal analysis, distinguishing where opacity, statistical inference and model governance create friction with legality and reviewability. Section V asks whether the AI Act amounts to a constitutional response in this sense, examining in particular the reliance on an internal-market legal basis, delegated norm-making, standardisation and codes of practice. Section VI then applies the Venice Commission's Checklist, element by element, legality, legal certainty, prevention of abuse of power, equality and non-discrimination, and access to justice, to evaluate the Act and related instruments; here access to justice functions as the decisive hinge, because without intelligible explanations, evidence preservation and disclosure, and contestation pathways, the rule of law cannot serve its instrumental role. Finally, Section VII sets out *de iure condendo*

²² Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, "2025 Rule of Law Report. The rule of law situation in the European Union", Strasbourg, 8.7.2025 COM(2025) 900 final.

²³ European Commission, "Methodology for the preparation of the Annual Rule of Law Report", 2025.

²⁴ The Venice Commission is a recurrent benchmark in the evaluation conducted in the Report. See, for instance, in the methodology where it is stated that the «Check List can help to identify specific risks and weaknesses», cit., 4.

proposals to re-align the framework with that role, including statutory bases for sensitive deployments, limits on soft-law reliance, clarified and enforceable oversight duties, independently verified equality safeguards, and remedies that guarantee reasons and evidence to courts and affected individuals. Throughout, the claim is: constitutionalisation is credible only if the rule of law works as a means. The contribution of the article is to provide a structured way to test that proposition in the AI domain and to indicate how the Union's framework can be re-oriented, doctrinally and practically, towards guarantees that are enforceable in fact, not merely in form.

2. Rule of Law: Theoretical Basis

In contemporary European constitutionalism, the rule of law is not a terminal value but a functional architecture. Its formal requirements have worth because they enable the substantive ones, equality, proportionality, and rights protection, to be exercised. On this view, the question to ask of any governance arrangement is not whether it can be described as “rule-of-law compliant” in the abstract, but whether it furnishes those who are governed with the concrete capacities to test power: a legal basis they can invoke, norms they can foresee, discretion they can challenge, and remedies they can obtain.

The rule of law has long occupied a central place in the European Union's constitutional imagination, but its precise content remains contested.²⁵ In Pitruzzella's formulation, its core elements are «the principle of legality, freedom and equality of citizens, separation of powers, independence of judges».²⁶ If the rule of law ensures that «it is the laws, not men, that govern us»,²⁷ and that the division of powers together with the principle of legality are the institutional devices that make this aspiration real.

As a matter of fact, the common law jurist Albert Venn Dicey famously defined the rule of law in three elements: the absolute supremacy of regular law as opposed to arbitrary power; equality before the law of all persons, regardless of status; and the constitution as the result of ordinary law, not a higher abstract text.²⁸ Dicey's view is often criticised as narrow and formalistic,²⁹ yet his insistence on equality before the ordinary courts still resonates in contemporary European jurisprudence. It anticipates one of the central doctrinal moves of the Court of Justice: that the availability of judicial review is an indispensable component of the Union's identity,

²⁵ For a brief overview, A. Z. Huq, *The Rule of Law: A Very Short Introduction*, Oxford, 2024.

²⁶ G. Pitruzzella, *Stato di diritto, indipendenza delle corti e sovranità popolare: armonia o conflitto*, in *Rivista di Diritti comparati*, 3, 2018, 12 ff.

²⁷ R. Bin, *Lo Stato di diritto*, cit., 3.

²⁸ A. V. Dicey, *Introduction to the Study of the Law of the Constitution*, London, 1959.

²⁹ M. D. Walters, *The Spirit of Legality: A. V. Dicey and the Rule of Law*, in J. Meierhenrich - M. Loughlin (eds.), *The Cambridge Companion to the Rule of Law*, Cambridge UK, 2021, 153 ff.

and that independence and impartiality are not political aspirations but legal requirements.

Modern legal philosophy has offered contrasting accounts of what the rule of law requires. Joseph Raz, in his seminal essay “The Rule of Law and Its Virtue,” insisted that the rule of law is a *formal* virtue of legal systems.³⁰ For Raz, law must be capable of guiding behaviour: it should be general, prospective, clear, relatively stable, and applied by an independent judiciary. This “thin” conception has been influential in both scholarship and jurisprudence and continues to underpin arguments that the rule of law is primarily about foreseeability and legal certainty.³¹

By contrast, Lon Fuller proposed that law possesses an «inner morality»³² consisting of requirements such as generality, publicity, clarity, non-contradiction, non-retroactivity, stability, and congruence between official action and declared rule. Although Fuller’s account is sometimes classified as formal, his insistence that these requirements express a moral dimension of legality suggests a bridge between the thin and thick conceptions:³³ even the formal demands of law presuppose a normative commitment to treating individuals as subjects capable of understanding and following rules. Fuller’s account resonates with the EU’s concern for legitimate expectations, non-retroactivity, and coherence, all of which are now entrenched as general principles of Union law.³⁴

Another relevant contribution comes from Jeremy Waldron’s emphasis on procedure, as anticipated, which illustrates why the rule of law cannot be reduced to abstract legality.³⁵ For him, the essence of legality lies in the possibility for individuals to contest decisions and to receive reasons: a view that resonates strongly with the Union’s insistence on effective judicial protection,³⁶ also addressed strongly by the Venice Commission’s understanding of the rule of law.³⁷ On the other hand, Paul Craig’s well-known distinction between formal and substantive conceptions confirms this point: rules may be general and clear, but without substantive guar-

³⁰ J. Raz, *The Rule of Law and its Virtue*, cit.

³¹ See P. Craig, *Formal and substantive conceptions of the rule of law: an analytical framework*, London, 2017.

³² L. L. Fuller, *The Morality of Law*, Yale, 1969.

³³ For a critical analysis of Fuller’s conceptualisation of «inner morality», see, *ex multis*, P. P. Nicholson, *The Internal Morality of Law: Fuller and His Critics*, in *Ethics*, 84(4), 1974, 307 ff.; K. Rundle, *The Morality of the Rule of Law: Lon L. Fuller*, in J. Meierhenrich - M. Loughlin (eds.), *The Cambridge Companion to the Rule of Law*, Cambridge UK, 2021, 186 ff.

³⁴ For a broader reflection on the topic, see T. Tridimas, *The General Principles of Law in the European Union Legal Order*, in T. Tridimas (eds.), *The General Principles of EU Law*, Oxford, 2006.

³⁵ J. Waldron, *The Rule of Law*, cit., 3.

³⁶ As argued by Pitruzzella, «the independence of judges falls within the scope of EU law and is considered essential to guarantee the fundamental right to a fair trial, which, in turn, is necessary to ensure all the rights that individuals derive from EU law, safeguarding the value of the rule of law», in G. Pitruzzella, *Stato di diritto, indipendenza delle corti e sovranità popolare: armonia o conflitto*, in *Rivista di Diritti comparati*, 3, 2018, 9 ff. (aut. tr.).

³⁷ S. Granata-Menghini-S. Ninatti, *Chapter 10: The evolving paradigm of human rights protection as interpreted and influenced by the Venice Commission*, in L. Violini-A. Baraggia (eds.), *The Fragmented Landscape of Fundamental Rights Protection in Europe*, Cheltenham, 2018, 201 ff.

antees of equality and rights, legality becomes hollow.³⁸ Conversely, rights guarantees without procedural channels degenerate into vague aspirations. András Sajó has warned that this fragility is intrinsic: the rule of law is frequently invoked rhetorically while eroded in practice.³⁹ This insight is particularly relevant to the Union, where rule-of-law discourse has become central in confronting constitutional backsliding,⁴⁰ but is equally applicable to the technological sphere. Artificial intelligence illustrates the same risk: legality may be formally present through compliance checklists, while, in practice, access to reasons and remedies is curtailed.⁴¹

The multiplicity of these accounts should therefore be read not as a weakness but as evidence of the layered nature of the rule of law in modern constitutionalism. Nonetheless, the fragility noted by Sajó, and proved in factual terms by the so-called “rule of law crisis” or the “rule of law backsliding”⁴² reminds us that these guarantees are never secured once and for all but require constant vigilance, especially when new modalities of exercising power, such as AI systems, enter the legal order.

In the European Union, this plural heritage has been crystallised into legal form. Art. 2 TEU elevates the rule of law to a foundational value;⁴³ art.

³⁸ P. Craig, *Formal and substantive conceptions of the rule of law: an analytical framework*, cit., 6.

³⁹ As the scholar acknowledges, «the rule of law is “an unqualified human good.” But the rule of law can be easily hijacked and turned into rule by law; it is also a technique of those who grabbed state power lawfully to impose their will in a seemingly legitimate manner. Moreover, it is often self-defeating or unable to provide guidance especially where political and politicized actors would like to abuse the legal loopholes»; in A. Sajó, *Militant Rule of Law*, in *Verfassungsblog*, 20 December 2023.

⁴⁰ C. I. Nagy, *The Diagonality Problem of EU Rule of Law and Human Rights: Proposal for an Incorporation à l'europpéenne*, in *German Law Journal*, 21(5), 2020, 838 ff.; O. Polański, *Poland: another episode of «rule of law backsliding» - Judgment P 7/20 and a threat to the integrity of the EU legal order*, in *Public Law*, 2022, 153 ff.; K. L. Scheppele et al., *EU Values Are Law, after All: Enforcing EU Values through Systemic Infringement Actions by the European Commission and the Member States of the European Union*, in *Yearbook of European Law*, 39, 2020, 3 ff.

⁴¹ As stressed from the digital point of view by G. De Gregorio-S. Demková, *The Constitutional Right to an Effective Remedy in the Digital Age: A Perspective from Europe*, in C. van Oirsouw et al. (eds.), *European Yearbook of Constitutional Law 2023: Constitutional Law in the Digital Era*, The Hague, 2024, 223 ff.

⁴² It goes beyond the scope of this paper to analyse the crucial and critical problem of the general erosion of the rule of law, since the focus rests upon its integration with the technology and, moreover, the AI perspective. However, on the notion of “rule of law backsliding” and the ongoing “rule of law crisis” in the EU, see *ex multis* N. Canzian, *The European Principle of Judicial Independence: The Case of Poland*, in *Quaderni costituzionali*, 2, 2020; A. von Bogdandy et al., *Un possibile «momento costituzionale» per lo Stato di diritto europeo: i confini invalicabili*, in *Quaderni costituzionali*, 4, 2018, 855 ff.; B. Bakó, *Hungary's Latest Experiences with Article 2 TEU: The Need for 'Informed' EU Sanctions*, in A. von Bogdandy et al. (eds.), *Defending Checks and Balances in EU Member States: Taking Stock of Europe's Actions*, Heidelberg, 2021, 35 ff.; M. Wiącek, *Constitutional Crisis in Poland 2015–2016 in the Light of the Rule of Law Principle*, in A. von Bogdandy et al. (eds.), *Defending Checks and Balances in EU Member States: Taking Stock of Europe's Actions*, Heidelberg, 2021, 15 ff.

⁴³ Nonetheless, as referenced by Schroeder, «the Commission and the courts when enforcing the rule of law vis-à-vis a certain Member State, however, cannot simply refer to art. 2 TEU and content themselves to an “I know it when I see it” approach». See W. Schroeder, *The Rule of Law As a Value in the Sense of Article 2 TEU: What Does It Mean and Imply?*, in A. von Bogdandy et al. (eds.), *Defending Checks and Balances in EU Member States: Taking Stock of Europe's Actions*, Heidelberg, 2021, 106.

19(1) TEU obliges Member States to ensure remedies sufficient to guarantee effective judicial protection in fields covered by Union law; and art. 47 of the Charter secures the individual right to «challenge in court the legality of any decision or any other national measure relating to the application of any Union act to them»⁴⁴ before an independent and impartial tribunal. The Court of Justice has treated these provisions as enforceable standards, most famously in *Associação Sindical dos Juizes Portugueses contra Tribunal de Contas* (C-64/16),⁴⁵ where it held that art. 19(1) TEU concretises the rule of law value in art. 2 and binds Member States whenever their courts may rule on EU law. In *Rosneft* (C-72/15),⁴⁶ the Court declared that «the very existence of effective judicial review designed to ensure compliance with provisions of EU law is of the essence of the rule of law». The lesson from these debates is that the rule of law is best understood not as a final state but as a dynamic principle that operates simultaneously at different levels. For European constitutionalism, it is also a structural principle that serves as the means for securing fundamental rights and constraining arbitrariness. This layered conception provides the bridge for analysing the impact of artificial intelligence. For if AI redistributes how power is exercised,⁴⁷ by embedding normative choices in technical design and by reconfiguring administrative and private decision-making, the rule of law is the measure by which we can determine whether these exercises of power remain legally constrained and rights protective.

2.1. The EU's Conception of Rule of Law: A Means, Not a Goal

The preceding discussion has shown that the rule of law cannot be reduced to a single, unitary formula. As described, it is a layered concept encompassing legality, certainty, equality, participation and access to judicial protection. Specifically, what emerges is that, if the rule of law functions as a means of ensuring that rights are real and arbitrariness constrained, the European Union provides a paradigmatic illustration of how that principle can be constitutionalised.⁴⁸ While art. 2 TEU elevates the rule of law to a foundational value, art. 19(1) TEU obliges Member States to provide remedies sufficient to secure effective judicial protection, and art. 47 of

⁴⁴ G. Pitruzzella, *Stato di diritto, indipendenza delle corti e sovranità popolare: armonia o conflitto*, cit., 6.

⁴⁵ See ECJ, Case C-64/16, *Associação Sindical dos Juizes Portugueses (ASJP) v Tribunal de Contas*, 27 February 2018, cit.

⁴⁶ See ECJ, Case C-72/15, *PJSC Rosneft Oil Company v Her Majesty's Treasury and Others*, 28 March 2017, §73.

⁴⁷ See F. Pasquale, *Inalienable Due Process in an Age of AI: Limiting the Contractual Creep toward Automated Adjudication*, in A. Reichman et al. (eds.), *Constitutional Challenges in the Algorithmic Society*, Cambridge UK, 2021, 42 ff.; S. Rosengrün, *Why AI is a Threat to the Rule of Law*, in *Digital Society*, 1, 2022, 2 ff.

⁴⁸ A. von Bogdandy et al., *Un possibile «momento costituzionale» per lo Stato di diritto europeo: i confini invalicabili*, cit.; G. Pitruzzella, *Stato di diritto, indipendenza delle corti e sovranità popolare: armonia o conflitto*, cit., 7.

the Charter enshrines an individual right to a remedy.⁴⁹ Taken together, these provisions show that formal legality and substantive protection are not alternatives. Formal legality without substantive protection would mean stability without justice; substantive rights without formal legality would mean aspirations without enforceability.⁵⁰

It is, however, the Court of Justice that has transformed the rule of law from a value to an enforceable standard by reading it through the combined lenses of art. 19(1) TEU and art. 47 CFR.⁵¹ Specifically, judicial independence, impartiality and access to remedies have been recognised as structural requirements of art. 19(1) TEU, indispensable for the right to effective judicial protection.⁵² In a series of cases against Poland,⁵³ the Court confirmed that measures undermining irremovability, equality between judges or the freedom to make preliminary references breach the rule of law. In *Repubblika*,⁵⁴ it made clear that these guarantees are not mere accession conditions but permanent obligations. Other decisions, such as *LM*,⁵⁵ linked the rule of law directly to individual rights. From this line of case law emerges a doctrinal triangle: art. 2 TEU formulates the value, art. 19(1) TEU imposes the obligation, and art. 47 CFR secures the individual right. Together, they constitutionalise the rule of law as both a structural principle of the Union and a justiciable entitlement for individuals.⁵⁶

⁴⁹ See, for instance, ECJ, Case C-362/14, *Schrems v. Data Protection Commissioner*, §60, «the European Union is a “union based on the rule of law”».

⁵⁰ Bast and von Bogdandy describe the rule of law as a “structural principle” of the Union, one that pervades the legal order and ensures the conditions under which rights can be vindicated. Their analysis makes clear that the rule of law has no independent existence apart from the guarantees it furnishes for legality, equality and judicial review. Similarly, Casolari has argued that the protection of fundamental rights in EU law is mediated through rule-of-law structures: without legality, clarity and review, rights remain aspirational. This underscores that the rule of law is not an ornament in the Union’s constitutional architecture but the mechanism through which values are translated into enforceable claims. See A. von Bogdandy *et al.*, *A Potential «Constitutional Moment» for the European Rule of Law: The Boundaries Not to Be Crossed*, in *Quaderni costituzionali*, cit., 7.

⁵¹ S. Peers *et al.*, *Article 47*, in *The EU Charter of Fundamental Rights: A Commentary*, Hart Publishing, 2021, 1245 ff.; K. Gutman, *The Essence of the Fundamental Right to an Effective Remedy and to a Fair Trial in the Case-Law of the Court of Justice of the European Union: The Best Is Yet to Come?*, in *German Law Journal*, 20(6), 2019, 884 ff.

⁵² S. Prechal, *Article 19 TEU and National Courts: A New Role for the Principle of Effective Judicial Protection?*, in G. Gentile *et al.* (eds.), *Article 47 of the EU Charter and Effective Judicial Protection, Volume 1: The Court of Justice’s Perspective*, United Kingdom, 2022.

⁵³ ECJ, C-619/18, *European Commission v Republic of Poland*, 24 June 2019 (on the independence of the Supreme Court); Case C-192/18, *European Commission v Republic of Poland* (on the independence of ordinary courts), 5 November 2019; Case C-791/19, *European Commission v Republic of Poland* (on the disciplinary regime for judges), 15 July 2021.

⁵⁴ In ECJ, Case C-896/19, *Repubblika v Il-Prim Ministru*, 20 April 2021, the Court stressed that judicial independence and effective remedies are not mere accession requirements but permanent obligations under art. 2 TEU.

⁵⁵ In Case C-216/18 PPU, *Minister for Justice and Equality v LM (Celmer)*, 25 July 2018, the Court further held that systemic deficiencies in judicial independence in a Member State may justify refusal to execute a European Arrest Warrant, thereby linking the rule of law directly to individual rights.

⁵⁶ As Smuha observes, rule-of-law discourse must incorporate human rights and

Ultimately, legislative and policy practices have reinforced this transformation. The Venice Commission's Rule of Law Checklist has provided a shared vocabulary for monitoring compliance,⁵⁷ while Regulation (EU) 2020/2092⁵⁸ and the Recovery and Resilience Facility have operationalised conditionality by linking access to EU funds to verifiable safeguards.⁵⁹ Most recently, the EU has embedded these benchmarks in its own institutional practice through the annual Rule of Law Report.⁶⁰ As clarified in the 2025 Communication on the methodology, the Report translates art. 2 TEU into concrete indicators covering legality, legal certainty, prevention of abuse of power, equality and access to justice. By structuring an annual cycle of monitoring around these benchmarks, the Commission has confirmed that the rule of law is conceived not as an aspiration but as a set of verifiable conditions for rights protection.⁶¹ This development aligns EU practice with the Venice Commission's Checklist, demonstrating a convergence between constitutional theory, judicial enforcement, and policy monitoring.⁶²

This evolution provides the benchmark for new regulatory domains. If that is the European Union's constitutional grammar, then any serious engagement with AI must show how the same conduits are maintained when decisions are generated by systems that learn, adapt and obscure their reasoning. That is the challenge of autonomy.

3. Rule of Law and AI: The Challenge of Autonomy

The preceding section reconstructed the Union's understanding of the rule of law as both structural and instrumental, anchored in legality, judicial independence, and effective remedies. That "constitutionalisation" provides the benchmark against which new forms of governance must be

accountability if it is to impose real constraints on emerging forms of power, in N. A. Smuha (ed.), *From Rule of Law to Algorithmic Rule by Law*, cit., 2.

⁵⁷ G. Repetto, *The constitutional relevance of the ECHR in domestic and European law*, Cambridge-Antwerp-Portland, 2013.

⁵⁸ Regulation (EU, Euratom) 2020/2092 of the European Parliament and of the Council of 16 December 2020 on a general regime of conditionality for the protection of the Union budget. See also *Rule of Law Conditionality Regulation: European Implementation Assessment | Think Tank | European Parliament*, 2025.

⁵⁹ D. Kochenov, *EU Enlargement and the Failure of Conditionality*, The Hague, 2008; for a critical standpoint, see A. Baraggia, *Democratic Blackmail. The Use of Conditionality to Protect the Rule of Law*, in *Quaderni costituzionali*, 2, 2023, 355 ff.

⁶⁰ Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, "2025 Rule of Law Report. The rule of law situation in the European Union", Strasbourg, 8.7.2025 COM(2025) 900 final.

⁶¹ As mentioned *supra*, a means more than a goal. See T. Tridimas, *The General Principles of Law in the European Union Legal Order*, cit., 6.

⁶² Not in contrast with the concept of autonomy of EU Law, as also reaffirmed in K. Lenaerts, *Upholding the Rule of Law through Judicial Dialogue*, in *Yearbook of European Law*, 38, 2019, 3 ff.

assessed.

Artificial Intelligence is the most immediate test of this standard: in particular, the emergence of AI autonomy marks a pivotal shift in the landscape of digital governance, with far-reaching implications for constitutional principles and regulatory frameworks. Consider automated decision-making systems in welfare or credit scoring:⁶³ when opaque algorithms determine access to essential services, individuals may be unable to understand or contest the reasons behind outcomes, undermining both autonomy and effective redress.⁶⁴ Even more directly, generative AI tools deployed in political campaigning illustrate how AI can manipulate individual preferences at scale, blurring the line between free choice and engineered behaviour.⁶⁵ These dynamics expose how AI risks converting the law's role from a framework for autonomous agency into a tool of subtle coercion.

By redistributing where and how decisions are made, AI systems bring into play precisely the guarantees the European Union has attached to the rule of law, as outlined *supra*.

As remarked also elsewhere, «the autonomy of AI systems challenges traditional notions of accountability and oversight, prompting scholars and policymakers to seek new frameworks that can properly regulate these technologies while upholding constitutional values such as democracy, the rule of law, and fundamental rights. For example, the lack of transparency of some AI decisions may pose a direct challenge to the constitutional principle of due process, which requires that decisions, particularly those affecting the rights of individuals, be understandable and appealable».⁶⁶

One way of understanding this evolution is through the different phases of digital constitutionalism.⁶⁷ Specifically, the first era revolved around addressing the power of platforms like Google, Facebook, and Amazon, ensuring they remained accountable within a framework of fundamental rights and democratic oversight. This era, often described as the first phase of was largely reactive, as courts like the ECJ played a pivotal role in translating traditional constitutional principles to the digital sphere.⁶⁸ However, as the algorithmic society gave way to the rise of AI-driven au-

⁶³ These are some of the most critical sectors which are indeed classified by the Regulation (EU) 2024/1689, as mostly “high risk”, therefore, subject to most of the obligations under the AI Act.

⁶⁴ T. Aimen, *Cognitive freedom and legal accountability: Rethinking the EU AI act's theoretical approach to manipulative AI as unacceptable risk*, in *Cambridge Forum on AI: Law and Governance*, 1, 2025, 20 ff.

⁶⁵ M. Groh *et al.*, *Human detection of political speech deepfakes across transcripts, audio, and video*, in *Nature Communications*, 15(1), 2024, 7629 ff.

⁶⁶ O. Pollicino, *Regolare l'Intelligenza Artificiale: la lunga via dei diritti fondamentali*, in O. Pollicino *et al.* (eds.), *La disciplina dell'intelligenza artificiale*, Milan, 2025. The quoted text was translated by the authors.

⁶⁷ For accounts that frame the evolution of digital constitutionalism in terms of successive phases, see O. Pollicino, *Potere digitale*, in M. Cartabia - M. Ruotolo (eds.), *Potere e Costituzione. Enciclopedia del diritto: i tematici*, Rome, 2023.

⁶⁸ The aspects hereby mentioned have been further analysed in O. Pollicino, *Judicial Protection of Fundamental Rights on the Internet: A Road Towards Digital Constitutionalism?*, London, 2021.

tonomy, the regulatory focus expanded.⁶⁹ No longer confined to holding platforms accountable, governance now faces the challenge of managing systems capable of self-learning, adaptation, and decision-making beyond explicit human instructions.

Consequently, a second phase, oriented towards the development of regulatory frameworks, started, e.g., with the General Data Protection Regulation (GDPR)⁷⁰: it was specifically designed to impose checks on the power of these platforms, ensuring that they did not violate individual rights.

Moreover, since the second half of the 2010s, the EU has adopted a series of sector-specific legislative measures aimed at introducing new obligations for providers – and especially for hosting service providers – in order to make them more responsible for combating illegal conduct and material online. Thus, with the Directive (EU) 2018/1808,⁷¹ the legislator amended the regulation of audiovisual media, introducing a series of obligations for providers of video-sharing platforms; the Directive (EU) 2019/790 innovated in the field of copyright to adapt the sector legislation to the new challenges of the Internet;⁷² the Regulation (EU) 2021/784 introduced a specific discipline aimed at combating online terrorist content.⁷³ Moreover, in 2022, with the approval of the Digital Services Act (DSA) a general and, so to speak, “horizontal” reform of the European regulatory framework on the moderation of online content was finally carried out.⁷⁴ Hence, the second phase represents a procedural turn, where the focus shifted towards creating robust mechanisms for ensuring transparency, accountability, and the enforcement of digital rights, by giving more and more supervisory powers to the European Commission.⁷⁵

This evolution represents a profound transformation in the regulatory landscape: one that demands a shift from principles of algorithmic governance to those capable of addressing the complexities of AI autonomy.

⁶⁹ O. Pollicino, *Potere digitale*, cit., 10.

⁷⁰ Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons concerning the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).

⁷¹ Directive (EU) 2018/1808 of the European Parliament and of the Council of 14 November 2018 amending Directive 2010/13/EU on the coordination of certain provisions laid down by law, regulation or administrative action in Member States concerning the provision of audiovisual media services (Audiovisual Media Services Directive) in view of changing market realities, 2018 O.J. (L 303) 69.

⁷² Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the digital single market and amending Directives 96/9/EC and 2001/29/EC.

⁷³ Regulation (EU) 2021/784 of the European Parliament and of the Council of 29 April 2021 on addressing the dissemination of terrorist content online, 2021 O.J. (L 172) 79.

⁷⁴ Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act), 2022 O.J. (L 277) 1.

⁷⁵ P. Van Cleynenbreugel, *The Commission's digital services and markets act proposals: First step towards tougher and more directly enforced EU rules?*, in *Maastricht Journal of European and Comparative Law*, 28(5), 2021, 667 ff.

Autonomous systems, unlike traditional automated processes, introduce novel challenges due to their ability to function without human oversight, relying on self-learning mechanisms and adaptive algorithms. While these capabilities enable unprecedented efficiency and innovation, they also pose risks to transparency, accountability, and the protection of fundamental rights. These challenges are particularly acute in areas like law enforcement, where AI's capacity for autonomous decision-making directly impacts privacy, non-discrimination, and other core rights.⁷⁶ The pervasive influence of the algorithmic society extends beyond information management into various other contexts, such as healthcare, labour, justice, and public security. In these sectors, too, we witness the widespread adoption of automated decision-making systems, which not only reinforce the role of algorithms but also highlight the increasingly intertwined relationship between public and private entities. This growing presence of private actors in traditionally public domains underscores a significant shift in governance structures.⁷⁷

However, as we transition from an algorithmic society to one increasingly driven by AI, this shift takes on even greater complexity. Moreover, the development of the AI regulatory framework discussed in this chapter was a direct response to the profound effects these systems were having on fundamental rights and societal norms. These regulations often emerged as reactive measures, spurred by public outrage over specific incidents where AI technology demonstrated its capacity to infringe on human rights.

A particularly prominent example is the use of COMPAS (Correctional Offender Management Profiling for Alternative Sanctions), an algorithm used in the U.S. criminal justice system to evaluate the likelihood of recidivism. This tool faced significant criticism for perpetuating racial bias, disproportionately labelling Black defendants as high-risk compared to their white counterparts.⁷⁸ Such bias undermined the fundamental rights to equality and a fair trial, as COMPAS-influenced decisions directly impacted sentencing and parole outcomes.⁷⁹ The opaque nature of the algorithm exacerbated the problem, as individuals had little insight into or recourse against the automated decisions.⁸⁰ This lack of transparency in decision-making processes posed serious concerns about due process and the right to a fair and impartial trial, two cornerstones of justice in democratic societies.⁸¹

⁷⁶ L. Violante, *Diritto e potere nell'era digitale. Cybersociety, cybercommunity, cyberstate, cyberspace: tredici tesi*, in *BioLaw Journal - Rivista di BioDiritto*, 1, 2022, 145 ff.

⁷⁷ The shift powerfully described by Balkin as the "Algorithmic Society", J. M. Balkin, *Free speech in the algorithmic society: Big data, private governance, and new school speech regulation*, in *UCDL rev.*, 51, 2017, 1149 ff.

⁷⁸ *State v. Loomis*, 881 N.W.2d 759 (Wis. 2016).

⁷⁹ J. Angwin-J. Larson-S. Mattu-L. Kirchner, *Machine Bias*, in *ProPublica*, 23 May 2016.

⁸⁰ F. Lagioia *et al.*, *Algorithmic fairness through group parities? The case of COMPAS-SAPMOC*, in *AI & Society*, 38(2), 2023, 459 ff.

⁸¹ As observed by Simoncini, the outcome of the US Supreme Court did not spark a violation on the right to "due process". Even though the use of these tools may, in fact, give rise to constitutionally relevant issues, but, in the present case, it reiterates the diriment argument of non-exclusivity: the algorithm was not the sole reason for the

Another high-profile incident was the Clearview AI scandal, which revealed that the company had scraped billions of images from social media without user consent to build a facial recognition database for law enforcement use.⁸² The widespread deployment of this technology, often without adequate public oversight, raised serious concerns about the right to privacy and data protection. Clearview AI's practices were particularly alarming as they infringed on individuals' control over their personal data, enabling unwarranted surveillance and tracking without their knowledge or consent.⁸³ This form of pervasive surveillance also had broader implications for freedom of expression and assembly, as individuals might hesitate to exercise these rights under the belief they were being constantly monitored. The case sparked international debates on the ethical use of facial recognition technology and highlighted its potential for misuse by both public and private entities.⁸⁴

In addition to these cases, the rise of generative AI has introduced new regulatory challenges, as demonstrated by the OpenAI case in Italy in 2023.⁸⁵ Italy's Supervisory Authority temporarily banned OpenAI's ChatGPT due to concerns over the unlawful collection and processing of personal data, and the lack of adequate mechanisms to prevent minors from accessing the platform. The case raised questions about how generative AI systems collect, store, and use personal information, especially when operating at such large scales.⁸⁶ The *Garante* ordered OpenAI to implement stronger privacy safeguards and transparency measures, including age verification tools and greater user control over their data. The incident attracted widespread attention, marking the first regulatory action taken against a major generative AI system in Europe. As a consequence of this huge societal echo, the regulation of such systems entered the radar of the European legislator, which eventually included specific provisions in the first European law on AI.⁸⁷

decision and, hence, was found to not being violating the principle of "due process". A. Simoncini, *L'algoritmo incostituzionale: intelligenza artificiale e il futuro delle libertà*, in *BioLaw Journal-Rivista di BioDiritto*, 15(1), 2019, 63 ff.

⁸² The case was addressed by numerous authorities: especially the Italian Supervisory Authority (SA, *Garante*) fined the US-based company Clearview AI EUR 20 million after finding it applied what amounted to biometric monitoring techniques also to individuals in the Italian territory, and it impeded the processing of biometric data of Italian citizens. Italian SA, Injunction Order against Clearview AI - 10 February 2022, Register of Measures, No. 50 of 10 February 2022. A similar direction was followed by the French Supervisory Authority (CNIL), Deliberation of the restricted committee N° SAN-2022-019 of 17 October 2022 concerning Clearview AI.

⁸³ The case and its far-reaching data protection and fundamental rights implications have also been analysed in this contribution:

⁸⁴ K. Hill, *The Secretive Company That Might End Privacy as We Know It*, in *The New York Times*, 18 January 2020.

⁸⁵ Italian SA, Measure of 30 March 2023, Register of Measures, No. 112 of 30 March 2023.

⁸⁶ G. Riva, *Diritto e Intelligenza Artificiale Generativa: L'istruttoria Del Garante per La Protezione Dei Dati Italiano Su OpenAI e ChatGPT*, in *Cyberspazio e diritto: rivista internazionale di informatica giuridica*, 25(1), 2024 153 ff.

⁸⁷ European Parliament, Amendments adopted by the European Parliament on 14 June

Henceforth, the episodes involving COMPAS, Clearview AI, and OpenAI highlight the urgent need for comprehensive AI regulations, particularly in light of the growing autonomy of these systems. AI's ability to function independently – learning, adapting, and making decisions with minimal human intervention – raises profound concerns about safeguarding fundamental rights. Issues such as discrimination, bias, privacy violations, and data security are exacerbated by the opacity and unpredictability inherent in autonomous AI systems. These cases reveal a pattern of reactive regulation, where lawmakers address specific incidents only after harm has occurred, rather than proactively anticipating the broader challenges posed by AI's autonomous capabilities. This reactive approach has created a precarious “cat-and-mouse” dynamic, with legislators scrambling to address the consequences of increasingly autonomous systems while struggling to develop long-term, comprehensive frameworks. As AI continues to evolve, this cycle of reactionary governance risks leaving significant gaps in oversight, exposing society to ongoing and unforeseen vulnerabilities. By failing to account for the unique challenges of AI autonomy, current regulatory efforts may fall short of effectively addressing the transformative and disruptive potential of these technologies.

4. Understanding Autonomy: The Technical Perspective

The growing autonomy of AI systems is indeed visible also from a technical perspective: AI's evolution from symbolic systems to machine learning and agentic architectures reshapes how power is exercised.⁸⁸ Early symbolic AI was rule-based: deductive instructions, written in logical languages, applied to data to produce transparent and predictable outputs.⁸⁹ Such systems posed legal questions, the validity of the legal basis, fairness of outcomes, but within relatively inspectable structures. Thus, contemporary AI is largely sub-symbolic.⁹⁰

2023 on the proposal for an Artificial Intelligence Act, COM(2021)0206 – C9-0146/2021 – 2021/0106(COD), Amendment 168.

⁸⁸ Transatlantic Reflection Group on Democracy and the Rule of Law in the Age of “Artificial Intelligence”, *A manifesto on enforcing law in the age of «artificial intelligence»*, in *European Law Journal: Review of European Law in Context*, 29(1–2), 2023, 249 ff.

⁸⁹ It is to note now that, then, the so-called intelligence of the system is associated with this ability to learn, and not just symbolise and, therefore, reproduce behaviours. This distinction is crucial to understanding the shift from automation and autonomy that is pursued by the AI Act. On this ability to infer and the elements characterising symbolic AI S. Harnad, *The Symbol Grounding Problem*, in *Physica D: Nonlinear Phenomena*, 42, 1–3, 1990, 335 ff.; S. Leonelli, *What distinguishes data from models?*, in *European Journal for Philosophy of Science*, 9(2), 2019, 22 ff.; S. Wachter-B. Mittelstadt, *A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI Survey: Privacy, Data, and Business*, in *Columbia Business Law Review*, 2, 2019, 494 ff. For a purely technical perspective, see M. L. Minsky, *Logical Versus Analogical or Symbolic Versus Connectionist or Neat Versus Scruffy*, in *AI Magazine*, 12(2), 1991, 34 ff.

⁹⁰ A. Akbari, *Facial Recognition Technologies 101: Technical Insights*, in M. Zalnieriute-R. Matulionyte (eds), *The Cambridge Handbook of Facial Recognition in the Modern State*,

The modern AI techniques, such as sub-symbolic AI approaches, encompass a variety of processes.⁹¹ Machine learning, deep learning and neural networks no longer rely on preset rules but extract patterns from large datasets. These techniques identify correlations, infer rules, and generate behaviour that appears autonomous. Generative AI,⁹² building on deep neural architectures, produces novel content – i.e., text, images, code – by modelling linguistic and statistical patterns. The most recent development is agentic AI, designed to pursue complex goals autonomously using reinforcement learning and adaptive decision-making, with limited human input.⁹³

Moreover, machine learning can be “supervised”, or “unsupervised”. The supervised approach systems are trained on labelled datasets where inputs are paired with known outputs. For example, in facial recognition, the system might be trained using pairs of images labelled as “same person” or “different person” to learn how to compare facial features accurately.⁹⁴ In a context which is, instead, “unsupervised”, it is the system that should identify patterns or clusters within the data on its own. Saying it differently, such connections are made without receiving instructions or learning from external sources, as in the supervised learning environment. In the biometric identification context, the use of unsupervised learning might involve group analysis distinguishing people for biometric features – long neck, big mouth, small eyes – or emotions, like happiness or anger, without knowing these features before or the exact names for these emotions. Therefore, the human “gives” to the system the data, without labelling “what we are looking for”.

Finally, the third category is “reinforcement learning”. This approach trains systems by rewarding desired behaviours and penalising undesired ones: «[it] is used when neither the samples nor the answers are available as a batch in the beginning».⁹⁵ These systems are, for instance, used for playing video games by maximising their score through trial and error. Another example might be that of self-driving cars. It starts with no knowledge of how to drive, but it learns by trial and error. The system is rewarded for actions like staying in the correct lane, stopping at traffic lights, and avoiding collisions, while penalties are applied for running red lights or hitting obstacles.

Cambirdge MA, 2024.

⁹¹ Among others, P. Wang-E. Fan, *Comparative Analysis of Image Classification Algorithms Based on Traditional Machine Learning and Deep Learning*, 141, 2021, 61 ff.

⁹² Art. 3(63) AI Act defines a “general-purpose AI model” as «an AI model, including where such an AI model is trained with a large amount of data using self-supervision at scale, that displays significant generality and is capable of competently performing a wide range of distinct tasks regardless of the way the model is placed on the market and that can be integrated into a variety of downstream systems or applications, except AI models that are used for research, development or prototyping activities before they are placed on the market».

⁹³ P. Bornet-J. Wirtz, *Agentic artificial intelligence: harnessing AI agents to reinvent business, work, and life*, Ebook, 2025.

⁹⁴ *Ibid.*

⁹⁵ A. Akbari, *Facial Recognition Technologies 101: Technical Insights*, cit., 14.

A particular branch of machine learning is that of “deep learning”.⁹⁶ This latter one is a specialised subset that uses neural networks to model complex, high-dimensional data. One of the most renowned uses of deep learning systems is specifically to create systems that can understand and analyse human language. Those systems present numerous and different applications, for instance, in the content moderation sector⁹⁷, as well as potential challenges due to their complexity and, *ça va sans dire*, opacity.

Key architectures in deep learning include: i) Convolutional Neural Networks (CNNs), ideal for image processing tasks, CNNs are widely used in facial recognition and medical imaging; Recurrent Neural Networks (RNNs), designed for sequential data, RNNs excel in tasks like speech recognition and time-series analysis; iii) Transformer Models, these are at the heart of large language models (LLMs) like OpenAI’s GPT-4, enabling advanced natural language processing and generation. In this sense, deep learning has transformed AI by enabling systems to tackle tasks that were previously thought to be impossible, such as real-time language translation and autonomous driving. However, it also introduces challenges. Neural networks often function as “black boxes,” making it difficult to interpret how they arrive at their decisions.

These technical shifts have direct constitutional implications. Symbolic AI challenged legality but remained tethered to explicit rules. Sub-symbolic and agentic AI redistribute decision-making into opaque technical architectures, diminishing foreseeability and accountability. When systems “learn” and adapt beyond their initial programming, it becomes harder to attribute responsibility, human oversight is reduced to auditing, and contestability is weakened. Autonomy in AI thus produces new rule-of-law concerns: ensuring legality, foreseeability, and effective remedies when decisions are no longer authored but generated.

Therefore, while early algorithmic systems primarily served to assist human decision-making by analysing vast amounts of data,⁹⁸ AI-driven systems now possess the capability to make decisions autonomously, for their technical advances described *supra*.⁹⁹

This change is more than a technical enhancement; it alters the dynamics of governance, as AI assumes responsibilities traditionally held by hu-

⁹⁶ «Deep-learning methods are representation-learning methods with multiple levels of representation at one level (starting with the raw input) into a representation at a higher, slightly more abstract level. With the composition of enough such transformations, very complex functions can be learned. For classification tasks, higher layers of representation amplify aspects of the input that are important for discrimination and suppress irrelevant variations [...] The key aspect of deep learning is that these layers of features are not designed by human engineers; they are learned from data using a general-purpose learning procedure», Y. LeCun *et al.*, *Deep learning*, in *Nature*, 521(7553), 2015, 436 ff.

⁹⁷ O. Pollicino-P. Dunn, *LA e valori democratici*, in *Intelligenza Artificiale e Disinformazione*, Milan, 2024.

⁹⁸ K. Yeung, *Algorithmic regulation: A critical interrogation*, in *Regulation & Governance*, 12(4), 2018, 505 ff.

⁹⁹ There are many areas in which AI is used and which affect the fundamental rights of individuals. Among others, it is to mention the law enforcement sector S. van Schendel, *Regulating Risk Profiling by Law Enforcement: A Task for Data Protection Law, Non-Discrimination Law and Criminal Procedural Law*, Tilburg, 2024.

mans, including tasks that directly affect constitutional rights such as freedom of expression, privacy and data protection.

Moreover, this change is reflected by the definition that the AI Act adopted for AI systems¹⁰⁰, AI as «a machine-based system designed to operate with varying levels of autonomy.»¹⁰¹ In this context, autonomy refers to the ability of AI systems to act independently of human intervention, making decisions or taking actions without requiring constant oversight.¹⁰² Adaptiveness, on the other hand, highlights the self-learning capabilities of these systems, allowing them to evolve, improve, or modify their behaviour over time as they interact with data and environments.

However, what truly distinguishes AI from conventional software is its ability to “infer”, as also mentioned by the very definition of AI systems in Regulation (EU) 2024/1689. While traditional software operates through predefined, step-by-step instructions set by developers, AI systems possess the capacity to analyse data, draw conclusions, and make decisions without explicit human direction. This ability to infer empowers AI systems to engage in higher-order tasks like learning, reasoning, and modelling, moving far beyond basic data processing.¹⁰³

This perspective underscores the complexity of AI systems compared to traditional software. AI’s ability to continuously learn, reason, and improve makes it a transformative technology that requires specific regulatory oversight.¹⁰⁴

Such an evolution – that in technical terms marked the passage from symbolic to sub-symbolic AI – is specifically characterised by the shift from algorithmic governance to AI-driven one: this passage represents a profound transformation of the regulatory landscape.¹⁰⁵

As a matter of fact, the term “algorithmic governance” traditionally refers to the use of automated decision-making (ADM) systems that process

¹⁰⁰ It is to note that the AI Act does not define AI, but just AI systems.

¹⁰¹ AI Act, art. 3, para. 1, no. 1). As mentioned *supra*, this definition is borrowed by the OECD.

¹⁰² AI Act, Recital 12.

¹⁰³ AI Act, Recital 12, «include machine learning approaches that learn from data how to achieve certain objectives; and logic- and knowledge-based approaches that infer from encoded knowledge or symbolic representation of the task to be solved», with either implementation going «beyond basic data processing» and enabling «learning, reasoning or modelling». Hence, the AI Act further elaborates on the techniques that enable this inference capability. Two main approaches are identified: machine learning methods, which allow systems to learn from data and adapt to achieve certain objectives, and logic- or knowledge-based approaches, which use encoded rules or symbolic reasoning to solve tasks. Both approaches enable AI systems to function in a way that is not only independent but also dynamic, adapting to new data and situations.

¹⁰⁴ As autonomy increases, assigning responsibility becomes gradually difficult, given the diminished importance or relevance or «control over the system due to its high degree of autonomy and capacity for self-learning», cit. A. Jedličková, *Ethical Approaches in Designing Autonomous and Intelligent Systems: A Comprehensive Survey towards Responsible Development*, in *AI & Society*, 2024.

¹⁰⁵ «Algorithmic technologies are not only instruments to exercise powers, which can interfere with fundamental rights, but can also be considered as rule-makers», in G. De Gregorio, *The Normative Power of Artificial Intelligence*, 2023, cit., 2.

large volumes of data based on pre-defined criteria.¹⁰⁶ ADM, in this context, involves using algorithms, a sequence of steps expressed in mathematical or logical terms, embedded within decision-making processes.¹⁰⁷ These systems range from fully automated decisions to hybrid approaches – i.e., purely automated decisions, automated triage, automated suspects and automated trials¹⁰⁸ – where algorithmic outputs assist human decision-making.¹⁰⁹ However, a recurring challenge lies in ensuring that these outputs do not merely shift responsibility away from human actors or obscure accountability through their complexity and opacity.¹¹⁰ While ADM systems have traditionally involved structured, rule-based automation, the emergence of AI introduces a layer of autonomy and adaptability, enabling systems to learn and evolve beyond static criteria. Unlike traditional algorithms, AI systems leverage advanced methodologies, such as machine learning, to generate outputs that are not always easily explainable, raising significant concerns about transparency, fairness, and accountability.¹¹¹ This shift from structured automation to autonomy represents a profound challenge for constitutional law, which has not yet fully adapted to address the opacity, adaptability and unpredictability of these systems.¹¹² Moreover, as Demková highlights, the right to human oversight and the ability to contest algorithmic decisions are safeguards that must be robustly implemented to mitigate the risks associated with ADM systems and uphold fundamental rights.¹¹³ Especially, as also underscored by Pasquale, the risk of the rising AI autonomy is that of eroding individuals,

¹⁰⁶ S. Demková, *Automated Decision-making and Effective Remedies: The New Dynamics in the Protection of EU Fundamental Rights in the Area of Freedom, Security and Justice*, Cheltenham, 2023; S. Lomborg et al., *Automated decision-making: Toward a people-centred approach*, in *Sociology Compass*, 17(8), 2023, 13097 ff.

¹⁰⁷ M. Busuioc, *Accountable Artificial Intelligence: Holding Algorithms to Account*, in *Public Administration Review*, 81, 2020.

¹⁰⁸ According to scholarly analysis, ADM can be classified as: solely automated decisions, automated triage, automated suspicion, and automated evidence. For a taxonomy of ADM see F. Palmiotto, *When Is a Decision Automated? A Taxonomy for a Fundamental Rights Analysis*, in *German Law Journal*, 25(2), 2024, 210 ff.

¹⁰⁹ Explanation and reasoning are the key aspects that, as a matter of fact, trigger the put into place of these technologies in the public actors' duties. On the duty of reasons of administrative law, see M. Fink-M. Finck, *Reasoned A(I)dmistration: explanation requirements in EU law and the automation of public administration*, in *European Law Review*, 47(3), 2022, 376 ff.

¹¹⁰ It is possible to distinguish between outputs. The classification is based on the position of the "data subject" and its relationship to the output of ADM. The first category of possible outputs of ADM includes evaluations, rankings, predictions, recommendations or classification of content. The ADM lists, classifies, ranks, filters, ranks or recommends. The second category of output concerns decisions made by the WMD concerning the person concerned. In these cases, the person interacting with the ADM is directly affected.

¹¹¹ L. Edwards-M. Veale, *Slave to the Algorithm? Why a «right to an explanation» is probably not the remedy you are looking for*, in *Duke Law & Technology Review*, 16, 2017, 18 ff.

¹¹² For a technical point of view, M. Soori et al., *Artificial intelligence, machine learning and deep learning in advanced robotics, a review*, in *Cognitive Robotics*, 3, 2023, 54 ff.

¹¹³ S. Demková, *Automated Decision-Making and Effective Remedies: The New Dynamics in the Protection of EU Fundamental Rights in the Area of Freedom, Security and Justice*, cit., 16.

trivialising concepts, like the rule of law, «a system of social relationships and legitimate governance, not simply the transfer and evaluation of information about behaviour».¹¹⁴

From a technical point of view, the opacity of ADM systems, often described as a “black-box”, *supra*, complicates efforts to ensure transparency and accountability, both of which are foundational to constitutional governance. In turn, in the legal environment, it has been advocated for a “right to explanation” of the automated decision, based on an interpretation of art. 22 of the GDPR. While this provision seeks to provide transparency by requiring meaningful information about the logic, significance and consequences of ADM usage, however, it has been criticised for its vagueness and limited practical applicability.¹¹⁵ This complexity, paired with the risk of superficial compliance by organisations, underscores the inadequacy of current mechanisms to ensure accountability.

Recently, the legal debate has been coupled also with the technical one, resulting in a shift from concepts like “explainable AI” or xAI:¹¹⁶ since the explanations, while necessary, do not always lead to genuine understanding, it is urged, instead, to strive for “understandable AI” (uAI), meaning going beyond mere explanation to focus on fostering understanding.¹¹⁷ Particularly, from a technical perspective, this shift can be achieved by building a framework that includes incoming propositions, knowledge bases, knowledge fusion, and confidence-change mechanisms to foster genuine understanding. This system will ensure that explanations are aligned with the users’ existing knowledge base and cognitive processes.¹¹⁸ Saying it in simple words, on the one hand, the purpose of xAI is to explain the internal workings of the model — for instance, by breaking down decisions into understandable components, creating *post-hoc* explanations, or simplifying

¹¹⁴ F. Pasquale, *A Rule of Persons, Not Machines: The Limits of Legal Automation*, in *George Washington Law Review*, 87(1), 2019, 1.

¹¹⁵ Critics argue that art. 22 does not explicitly mandate full explanations but only “meaningful information,” leaving ambiguity about what level of transparency is sufficient. Furthermore, ADM systems, particularly those based on machine learning, make it difficult to offer clear, actionable insights into their decision-making processes. L. Edwards-M. Veale, *Slave to the Algorithm? Why a «right to an explanation» is probably not the remedy you are looking for* (preprint), cit.; M. Fink-M. Finck, *Reasoned A(I)dmistration: explanation requirements in EU law and the automation of public administration*, cit.; S. Wachter *et al.*, *Why a right to explanation of automated decision-making does not exist in the General Data Protection Regulation*, in *International Data Privacy Law*, 7(2), 2017.

¹¹⁶ Explainable AI refers to methods and techniques in artificial intelligence (AI) that make the functioning of AI models understandable to humans. The primary objective of xAI is to ensure that the decisions and actions made by AI systems can be interpreted and explained, thereby fostering trust, transparency, and accountability. In particular, xAI aims to provide insights into the internal mechanisms of complex AI models, such as neural networks, which are often considered “black boxes” due to their opacity. See K. Grieman, *Explainable AI (XAI)*, in *Law, Death, and Robots: The Regulation of Artificial Intelligence in High-Risk Civil Applications*, London, 2024, 44 ff.

¹¹⁷ H. Abbass *et al.*, *Editorial: From Explainable Artificial Intelligence (xAI) to Understandable Artificial Intelligence (uAI)*, in *IEEE Transactions on Artificial Intelligence*, 5(9), 2024, 4310 ff.

¹¹⁸ V. Rodríguez-Doncel *et al.*, *Introduction: A Hybrid Regulatory Framework and Technical Architecture for a Human-Centered and Explainable AI*, in V. Rodríguez-Doncel *et al.* (eds.), *AI Approaches to the Complexity of Legal Systems XI-XII*, Cham, 2021, 1 ff.

the model to make it interpretable;¹¹⁹ on the other hand, understandable AI goes beyond merely explaining how decisions are made. It emphasizes creating AI systems that foster a deep understanding for users, especially non-experts.¹²⁰ uAI focuses on aligning the system's explanations with the user's cognitive processes and prior knowledge to ensure that users not only receive an explanation but also comprehend it fully. In a nutshell, uAI seeks to make AI models and their outputs more intuitive and meaningful to a wider range of users.

The implications from a are such that appeared in a recent ruling by the Constitutional Court of Colombia underscore the importance of human oversight in AI-driven decision-making. In this case, a judge used ChatGPT 3.5 to assist in drafting a legal decision.¹²¹ The court affirmed that while AI can complement human judgment, it must not replace it. The judge did not rely on AI to resolve the case but rather employed it as a tool to assist in writing the decision, based on reasoning provided by the judge. This case illustrates how AI can be integrated into decision-making in a way that respects human authority and maintains a clear boundary between assistance and autonomy.¹²²

Coming back to the European Union context, in the *Schufa* ruling by the ECJ, it emerged that what constitutes “automated decision-making” is not only the actual output of algorithmic analysis.¹²³ The ECJ established that even if a human is involved in the final decision-making process – e.g., a bank deciding whether to grant a loan – if an automated system, as a credit scoring algorithm, plays a key role in influencing that decision, it may still qualify as automated decision-making under data protection laws.¹²⁴ This ruling drew attention to a crucial matter: it is to be considered from a legal stand, not only the “algorithmic output”, but the result as a whole, even though a human is involved in the final stage.¹²⁵ Essentially, the court found that while the final decision to grant or deny credit may involve human intervention, the credit score itself plays a crucial role in that decision-making process. As such, the automated calculation of the credit score is an integral part of the overall automated decision-making process. Henceforth, the turn from algorithmic governance to AI governance rep-

¹¹⁹ H. Abbass *et al.*, *Editorial: From Explainable Artificial Intelligence (xAI) to Understandable Artificial Intelligence (uAI)*, cit., 17.

¹²⁰ B. D. Mittelstadt *et al.*, *The ethics of algorithms: Mapping the debate*, in *Big Data & Society*, 3(2), 2016, 679 ff.

¹²¹ República de Colombia Corte Constitucional, Ruling T-323/2024, in particular §§ 422-424.

¹²² *Ibid.*, § 424, lett. i).

¹²³ See Case C-634/21, *OQ v. Land Hesse*, 7 December 2023, cit.

¹²⁴ L. A. Bygrave, *Article 22 Automated individual decision-making, including profiling*, in C. Kuner *et al.* (eds.), *The EU General Data Protection Regulation (GDPR): A Commentary*, Oxford, 2020.

¹²⁵ A. Aza, *Scores as Decisions? Article 22 GDPR and the Judgment of the CJEU in SCHUEA Holding (Scoring) in the Labour Context*, in *Industrial Law Journal*, 2024, 35; G. Fasano, *L'interpretazione estensiva della nozione di “decisione automatizzata” ad opera della Corte di giustizia: una prospettiva più ampia ma ancora fragili tutele per le libertà fondamentali*, in *Rivista italiana di informatica e diritto*, 6(2), 2024, 15 ff.

resents a profound transformation in the regulatory landscape, where this shift is essentially driven by the increasing reliance on technological autonomy.¹²⁶ This transition, fueled by the increasing reliance on autonomous and adaptive AI systems, fundamentally alters the dynamics of accountability and oversight, necessitating a rethinking of regulatory structures. For the European Union, the central question is whether its regulatory architecture can translate the classical guarantees of the rule of law – legality, fairness, transparency, oversight and remedies – into operational duties fit for AI. The AI Act presents itself as the constitutional response: a framework intended to prohibit practices deemed incompatible with rights, to discipline high-risk systems, and to embed transparency, oversight and accountability into the AI lifecycle. The question is whether this response restores the conduits of the rule of law where autonomy obstructs them, or whether it codifies a compliance regime framed by market integration rather than by rights protection.

5. The AI Act: A Constitutional Response?

The AI Act is a landmark piece of legislation by the European Union, designed to create a “new” legal framework for AI technologies. As a matter of fact, the AI Act promise relied on the idea of building a legal framework of «rules for AI available in the Union market or otherwise affecting people in the Union [that] should therefore be human-centric, so that people can trust that the technology is used in a way that is safe and compliant with the law, including the respect of fundamental rights».¹²⁷

The Regulation, even though of pioneering importance, is not the only attempt at regulating AI. Almost contemporary to the EU, the Council of Europe worked and published its Framework Convention on Artificial Intelligence,¹²⁸ establishing rules applicable only to public actors, being the first binding treaty on AI, focusing on protecting human rights, democracy and the rule of law.¹²⁹ The growing momentum in regulating AI reflects the increasing global recognition of the profound impact that AI technologies can have on society, necessitating comprehensive legal frameworks to address the ethical, social, and legal challenges posed by these rapidly advancing systems.

¹²⁶ As observed by De Gregorio, «algorithmic technologies are not only instruments to exercise powers, which can interfere with fundamental rights, but can also be considered as rule-makers», in G. De Gregorio, *The Normative Power of Artificial Intelligence*, cit.

¹²⁷ Explanatory Memorandum published within the proposal for a Regulation of the European Parliament and of the Council, laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain union legislative acts, Brussels, 21.4.2021, COM(2021) 206 final.

¹²⁸ Council of Europe (CoE), Framework Convention on artificial intelligence, human rights, democracy and the rule of law, published on 17 May 2024, and open for signatures since 5 September 2024, Council of Europe Treaty Series - No. 225.

¹²⁹ The Framework Convention has a broader scope of application than the AI Act since it was negotiated by the COE’s Member States, together with Ministers or high-level representatives from the Governments of the United States, Canada, Mexico, Japan, Israel, Ecuador, Peru, Uruguay, and Argentina.

On the one hand, the regulatory push highlights the urgent need for international cooperation and alignment to ensure that AI development and deployment are guided by common ethical and legal principles. On the other hand, with the simultaneous efforts by various international bodies, it will be more complex for the European Union to establish undisputed leadership in AI regulation, as it previously did with the General Data Protection Regulation (GDPR)¹³⁰ and, to a lesser extent, the Digital Services Act (DSA).¹³¹ The shared efforts in the regulation of AI may dilute the EU's capacity to dominate the conversation entirely, making it essential for the EU to maintain a competitive edge in shaping global AI governance.¹³² Nonetheless, keeping with the tradition already established by other EU regulations, especially the GDPR and the DSA,¹³³ the AI Act is rooted in a risk-based approach, categorising AI systems into different tiers based on the level of risk they pose, ranging from minimal to prohibited.¹³⁴ This approach is intended to safeguard fundamental rights, promote transparency, and build trust in AI systems while maintaining a competitive and harmonised single market within the EU.¹³⁵

The AI systems are categorised into four tiers. At the top of the hierarchy are AI systems that are deemed to have such severe impacts on individual rights that they are outright banned. These include for instance, systems of biometric recognition used in real-time in public spaces for purposes of identification by law enforcement authorities,¹³⁶ social scoring systems, or those that use subliminal techniques to manipulate individuals or groups in ways that may cause them harm.¹³⁷

The second level is occupied by high-risk systems.¹³⁸ This category covers systems that, while not prohibited, pose significant risk to safety, fundamental rights, and societal interests. The AI Act specifically outlines high-risk sectors in Annex III, including areas such as biometric identification, critical infrastructure, education and vocational training, employment and worker management, access to essential private and public services, law

¹³⁰ Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of individuals with regard to the processing of personal data and on the free movement of such data and repealing Directive 95/46/EC (General Data Protection Regulation).

¹³¹ Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a single market for digital services and amending Directive 2000/31/EC (Digital Services Act).

¹³² F.P. Levantino-F. Paolucci, *Advancing the Protection of Fundamental Rights Through AI Regulation: How the EU and the Council of Europe are Shaping the Future*, in *European Yearbook on Human Rights*, Leiden, 2025.

¹³³ P. Dunn-G. D. Gregorio, *The European risk-based approaches: Connecting constitutional dots in the digital age*, in *Common Market Law Review*, 59(2), 2022.

¹³⁴ In particular, Annex III further specifies art. 6 of the AI Act and lists the AI systems which are considered “by default” by the EU legislator “high risk”.

¹³⁵ AI Act, art. 3, para. 1, no. 2), «risk means the combination of the probability of an occurrence of harm and the severity of that harm».

¹³⁶ AI Act, Annex II details the crime for which such AI systems can be put into place.

¹³⁷ AI Act, art. 5.

¹³⁸ AI Act, art. 6 and Annex III.

enforcement, migration, asylum, and border control, and judicial administration. This is the most comprehensive category in the AI Act, with the majority of the regulation dedicated to ensuring that these systems are transparent, fair and accurate.¹³⁹

The third level is occupied with limited-risk AI systems.¹⁴⁰ These systems are designed for human interaction, the creation or manipulation of content (including deepfakes),¹⁴¹ and image or audio generation systems. While not classified as high-risk, they still pose certain challenges, particularly regarding transparency. As a result, they are subject to specific transparency obligations, ensuring that users are aware when interacting with AI systems.¹⁴²

At the base of the hierarchy are AI systems with minimal risk, which are largely exempt from stringent requirements. Providers of these systems are encouraged to adhere to voluntary codes of conduct and governance mechanisms, promoting self-regulation.¹⁴³

Typically, the risk-based approach is illustrated as a pyramid, with minimal risks at the base and progressively greater risks as you move upward. However, the AI Act can be more aptly compared to an elevator than a pyramid for two key reasons.

First, it is up to the providers themselves to document and determine whether a system does not qualify as high-risk under a self-pre-assessment of the system.¹⁴⁴ By the way, this specific aspect contains potential loopholes regarding the consistency and robustness of compliance, as it places a considerable burden on providers to police themselves potential conflicts of interest.¹⁴⁵ Second, certain obligations designed for lower-risk systems also extend to higher-risk categories. A clear example of this is art. 50's transparency requirement, which applies not only to systems classified as limited risk but also to high-risk systems like those used for emotion recognition.¹⁴⁶

The most striking illustration of this “elevator mechanism” can be seen in the regulation of General Purpose AI (GenAI) systems. These systems follow a tiered structure. Defined as «an AI system based on a general-purpose AI model, capable of serving multiple purposes, both for direct use and for integration into other AI systems» (art. 3, no. 66), GenAI systems present unique regulatory challenges due to their broad applicability. These models are subject to a risk-based approach, with systems classified as having systemic risks facing stricter regulation compared to those with-

¹³⁹ AI Act, arts. 8-27.

¹⁴⁰ AI Act, art. 50.

¹⁴¹ AI Act, art. 50, para. 4.

¹⁴² AI Act, Recital 132.

¹⁴³ AI Act, Recital 165.

¹⁴⁴ AI Act, art. 6, para. 3.

¹⁴⁵ S. Wachter, *Limitations and loopholes in the EU AI Act and AI Liability Directives: what this means for the European Union, the United States, and beyond*, in *Yale Journal of Law and Technology*, 26(3), 2024.

¹⁴⁶ AI Act, art. 50, para. 3.

out systemic risks.¹⁴⁷

Although GenAI models are integral to these systems, they do not constitute AI systems on their own. The AI Act's provisions apply to GenAI models only when they are integrated into a functioning AI system. The obligations for GenAI model providers kick in once these models are placed on the market and continue even when integrated into an AI system. However, these obligations do not apply when the model is used solely for internal processes that do not affect individual rights or provide products or services to third parties. On the other hand, GenAI models with systemic risks are always subject to the AI Act's obligations due to their potentially significant adverse effects.¹⁴⁸

In essence, the AI Act treats GenAI systems as AI systems, and when a general-purpose AI model is integrated into an AI system, this integrated system is classified as a GenAI system if it gains the ability to serve a variety of purposes.¹⁴⁹ High-risk AI systems, including high-risk GenAI systems, are subject to extensive obligations under the AI Act, such as conducting a Fundamental Rights Impact Assessment (FRIA).¹⁵⁰ This assessment requires deployers to consider the protection of fundamental rights, thereby compelling both public and private actors to evaluate the potential impact of AI systems on these rights.

Nonetheless, this regulatory architecture poses significant challenges for deployers, who may face difficulties in ensuring comprehensive compliance. The complexity of evaluating AI's impact on diverse fundamental rights, coupled with the flexibility of GenAI systems to serve various functions, can make it hard for actors to anticipate all potential risks and to tailor safeguards accordingly. Moreover, the reliance on self-assessment increases the risk of underreporting or misjudging these impacts, especially when there is limited external oversight to ensure that the FRIA process is thoroughly and objectively conducted.¹⁵¹ This raises concerns about whether AI systems, particularly those classified as high-risk, will be deployed with adequate safeguards in place to prevent harm to individuals and society.

Moreover, different from the GDPR and the DSA, which strongly focus on the protection of a couple of fundamental rights mainly – privacy and

¹⁴⁷ The classification for systemic risk is further expanded in the guidelines of the “Guidelines on the scope of the obligations for general-purpose AI models established by Regulation (EU) 2024/1689 (AI Act)”, published by the European Commission, 18 July 2025. This designation carries significant regulatory consequences, including stricter obligations, that must not be mistaken with prohibited or high risk AI systems, under art. 5 and 6 of the AI Act. While these latter ones are about the *use* of the *system*, the “systemic risk” notion applies, instead, on different criteria: i) the model displays capabilities matching or exceeding those of the most advanced models on the market; ii) the Commission designates the model as systemic risk based on impact criteria set out in Annex XIII of the AI Act.

¹⁴⁸ AI Act, Recital 112.

¹⁴⁹ AI Act, Recital 100.

¹⁵⁰ A. Consentini *et al.* *Assessing the Impact of Artificial Intelligence Systems on Fundamental Rights*, in *MediaLaws.eu*, 31 March 2025.

¹⁵¹ S. Wachter, *Limitations and loopholes in the EU AI Act and AI Liability Directives: what this means for the European Union, the United States, and beyond*, cit., 20.

data protection, the former, freedom of expression, the latter –, the AI Act has a completely different scope of application. As it is evident even from the choice of the legal basis,¹⁵² the regulation on AI is a product safety norm that intersects with a wider range of issues.¹⁵³ However, when necessary, the AI Act also integrates the protection of other fundamental rights to address specific concerns, such as the inclusion of provisions related to data protection, especially in contexts where the processing of special categories of personal data is concerned, as is the case with biometric surveillance technologies.¹⁵⁴

The wide range of cases in which AI can trigger fundamental rights is especially evident in the case of the FRIA: this assessment is a rights-based one, meaning that the deployer is asked to consider the risks that an AI system may pose by pondering every right in the CFREU, especially those «in the areas of education, healthcare, social services, housing, administration of justice».¹⁵⁵ The impact assessment, much like the risk classification, suffers from the Act's over-reliance on self-assessment methodologies. Even though the deployer will have to communicate the results of the FRIA to the Market Surveillance Authority,¹⁵⁶ this approach risks becoming self-serving, as deployers may tailor the assessments to minimise perceived risks or liabilities, rather than conducting a thorough and impartial evaluation.¹⁵⁷ The lack of stringent external oversight exacerbates this issue, leaving room for biased or incomplete evaluations that could fail to adequately safeguard fundamental rights.

The AI Act, despite its pioneering ambition, exhibits an over-reliance on previous regulatory instruments, such as the GDPR, especially, and the DSA to some extent. While these frameworks have been effective within their respective domains, they are ill-equipped to address the distinctive challenges posed by the autonomy of AI systems. This over-reliance highlights the EU legislator's inadequate understanding of the unique risks associated with AI autonomy. As a matter of fact, in the AI Act, despite the promises,¹⁵⁸ the human element is minimally considered, largely relegated to the status of a potential “victim” of AI-induced harm,¹⁵⁹ rather than being envisioned as an integral actor in governance and oversight mechanisms.

This results in a regulatory framework where human involvement is reactive, appearing primarily in the context of mitigating risks rather than ensuring active and meaningful oversight of AI systems. Specifically, this

¹⁵² The AI Act is primarily grounded in art. 114 TFEU, which focuses on establishing and ensuring the functioning of the internal market.

¹⁵³ AI Act, Recitals 1 and 9.

¹⁵⁴ AI Act, Recitals 3 and 59.

¹⁵⁵ AI Act, art. 27, to be read together with art. 1, Recitals 93 and 96. The quote comes from Recital 96.

¹⁵⁶ *Ibid.*

¹⁵⁷ M. Ebers, *Truly Risk-Based Regulation of Artificial Intelligence - How to Implement the EU's AI Act*, SSRN Scholarly Paper, Rochester, NY, 2024.

¹⁵⁸ AI Act's Explanatory Memorandum, *ibid.*

¹⁵⁹ AI Act, art. 27, para. 1, lit. b) and c).

aspect is also reflected in the design of the remedies under the AI Act's arts. 85, 86 and 99, para. 10. Art. 85 enables individuals, or groups of individuals, to launch a complaint with a market surveillance authority «having grounds to consider that there has been an infringement of the provisions of this Regulation».¹⁶⁰ Art. 99, para. 10, recalls that access to effective judicial remedies and due process rights should be granted via procedural safeguards.¹⁶¹ Eventually, art. 86 also conceives a right for the individual to receive a «meaningful explanation of the role of the AI system in the decision-making procedure and the main elements of the decision taken».¹⁶² This newly introduced right to explanation attempts to address ongoing debates in the academic community,¹⁶³ particularly those surrounding the ambiguous nature of information rights under the GDPR. While the GDPR grants individuals access to information about the logic of algorithmic decisions under art. 22, the precise nature of these rights has been subject to interpretation. The AI Act now steps into this debate, providing a more explicit but still somewhat ambiguous formulation of the right to explanation. For example, the lack of clarity around how this right will apply in practice, especially concerning the role of AI-driven decisions in the broader decision-making process, remains a challenge. As this text affirms, these limitations underscore the need for clearer guidance on how AI systems should integrate human oversight and provide accountability in decision-making processes.

Hence, the human element is often relegated to mitigating harm only after it has occurred, rather than being proactively engaged in ensuring that AI systems operate within the bounds of fairness and legality from the outset. This reactive approach weakens the potential for robust due process protections, as it limits opportunities for individuals to challenge AI-driven decisions before they lead to detrimental consequences.

Moreover, the AI Act should not only ensure an effective and efficient right to remedies that allows individuals to contest the specific decisions made by an AI system but also extend accountability to the deployers of AI systems. This would involve holding the responsible entity accountable for its decision to implement AI in a particular context. In other words, procedural protections should go beyond merely challenging an AI decision and should encompass scrutiny of whether the decision to use AI in a given situation was appropriate and justified. This dual layer of responsibility – both for the decision and the deployment – would ensure a more comprehensive framework for safeguarding fundamental rights.

Besides, the AI Act fails to adequately account for the specific contexts in which AI technologies are deployed, leading to the potential for misinterpretation by Member States: this misalignment risks over-regulation in certain areas and insufficient attention in others, creating uneven en-

¹⁶⁰ *Ibid.*, art. 85.

¹⁶¹ Thus, Recital 27 specifies that «Without prejudice to the legally binding requirements of this Regulation and any other applicable Union law, those guidelines contribute to the design of coherent, trustworthy and human-centric AI, in line with the Charter and with the values on which the Union is founded».

¹⁶² *Ibid.*

¹⁶³ The referral is also to para. 2.1 of the present paper.

forcement and regulatory fragmentation across the Union.¹⁶⁴ One of the most glaring oversights is the Act's deliberate exclusion of AI in military applications from its regulatory scope.¹⁶⁵ By neglecting to address this crucial domain, the EU has left a significant gap in oversight, allowing the deployment of military AI technologies to proceed without the rigorous safeguards applied to civilian sectors.

The next section will, then, further unpack these critical aspects, trying to figure out what of the current system risk to harness the rule of law using as a litmus test the conceptualisation of the Venice Commission.

6. Testing the Rule of law in Context: Drawing from the Venice Commission

Having set out the general architecture of the AI Act, the crucial step is to test its provisions against the Union's own understanding of the rule of law. As argued in paragraph 2, the rule of law in the EU is not a lofty aspiration but a *functional principle*: a means of ensuring that fundamental rights are guaranteed through legality, foreseeability, accountability and remedies. In this sense, besides the scholarship analysis of the rule of law, a concrete benchmark arrives from the Venice Commission.¹⁶⁶ In this context, the Venice Commission also adopted a Checklist as «a tool for assessing the Rule of Law in a given country from the viewpoint of its constitutional and legal structures, the legislation in force and the existing case-law. [...] The checklist is mainly directed at assessing legal safeguards [...] The checklist is meant as a tool for a variety of actors who may decide to carry out such an assessment».¹⁶⁷ However, as the Explanatory Note recalls, the Checklist does not seek to redefine the concept, but to «make it justiciable and assessable», by breaking down the rule of law into concrete benchmarks that can be tested across legal systems. The goal is to move from the invocation of the rule of law as a political aspiration to its evaluation as a functional guarantee.

The Commission identified five clusters of requirements that together give content to the rule of law: legality, legal certainty, prevention of abuse of power, equality before the law and non-discrimination, and access to justice before independent courts. These criteria are not random: they correspond to the elements without which the rule of law ceases to operate as a *means* of protecting rights. As the next paragraph will develop, legality ensures that power is exercised through properly enacted law, rather than discretion or administrative practice; legal certainty makes the law

¹⁶⁴ For instance, Italy adopted the law 132/2025 for the implementation of the AI Act, that includes a modification of the criminal code to include art. 612-quater, "Unlawful dissemination of content generated or manipulated by artificial intelligence systems", which provides for a penalty of one to five years' imprisonment.

¹⁶⁵ AI Act, art. 2, para. 3.

¹⁶⁶ "Report on the Rule of law", adopted in Venice, March 25-26, 2011 - CDL-AD (2011)003rev., and the "Rule of Law Checklist", adopted in Venice, March 11-12, 2016 - CDL-AD(2016)007.

¹⁶⁷ *Ibid.*, paras. 23-28.

foreseeable and knowable, allowing individuals to orient their conduct; prevention of abuse of power guards against arbitrariness by requiring that discretion be constrained and accountable. Finally, equality secures fairness and prevents exclusionary or discriminatory outcomes, and access to justice guarantees that individuals have remedies before independent courts, without which all other principles remain aspirational.

The impact of the Checklist is not just limited to the CoE domain, since it has become a reference in European constitutional discourse because it translates a principle of high abstraction into criteria that can be applied to concrete cases.¹⁶⁸ Thus, it fits the approach of this paper: treating the rule of law not as a symbolic goal, but as a practical yardstick to test whether regulatory frameworks, such as the AI Act, actually secure remedies.¹⁶⁹ As mentioned *supra*, the Venice Commission Checklist is also referenced by the 2025 Rule of Law Report methodology, which illustrates how these benchmarks are now entrenched in Union practice.¹⁷⁰ Each of the five dimensions, legality, certainty, abuse of power, equality, and access to justice, is compatible with the concrete indicators for Member State assessment used by the European Commission. This convergence between the Venice Commission and the Commission's own reporting architecture underscores the appropriateness of using the same framework to evaluate the AI Act. Just as the Report measures whether Member States provide the institutional conditions for the rule of law to function, the question here is whether the regulatory instrument for Artificial Intelligence, the AI Act, embeds the operational guarantees that make remedies effective in the algorithmic domain.

For all of these reasons, the analysis that follows applies these five dimensions to the AI Act in order to find potential answers to whether, element by element, the Regulation meets the functional standard as inspired by the Venice Commission's understanding of the rule of law, or whether it remains confined to formal legality without delivering the substantive guarantees of the rule of law.

¹⁶⁸ *Council conclusions on fundamental rights and rule of law and on the Commission 2012 Report on the Application of the Charter of Fundamental Rights of the European Union*, Justice and Home Affairs Council Meeting, Luxembourg, 6-7 June 2013, part c, available at consilium.europa.eu. Moreover, in the "Resolution of the Italian Parliamentary Assembly", no. 2187/2017, it is specified that the Venice Commissione «invite international and regional organisations, including the Council of Europe as a whole and the European Union, to refer regularly to the Rule of Law Checklist in their relevant work. In this connection, the Assembly congratulates the Secretary General of the Council of Europe on having taken the Rule of Law Checklist into account in his 2017 annual report on the situation of democracy, human rights and the rule of law in Europe, and urges him to do so systematically in all his future annual reports», §6.4.

¹⁶⁹ On the relationship between the EU and the CoE on the regulation of AI, see F. P. Levantino-F. Paolucci, *Chapter 1 Advancing the Protection of Fundamental Rights through AI Regulation*, cit., 19.

¹⁷⁰ Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, "2025 Rule of Law Report. The rule of law situation in the European Union", Strasbourg, 8.7.2025 COM(2025) 900 final; European Commission, "Methodology for the preparation of the Annual Rule of Law Report", 2025.

6.1. Legality

The first benchmark of the Checklist is legality, understood as the requirement that all public power must be exercised on the basis of clear, accessible and foreseeable law, properly enacted by parliament and subject to review. Applied to the AI Act, legality raises two questions: the adequacy of the Act's own legal basis at the Union level,¹⁷¹ and the sufficiency of legal bases for the deployment of AI in Member States.¹⁷²

At the Union level, the AI Act rests exclusively on art. 114 TFEU, the internal market harmonisation clause. Formally, this satisfies legality: the EU acted within a recognised competence, the Act was adopted by the ordinary legislative procedure, and it takes the form of a binding regulation. Yet the reliance on art. 114 is constitutionally ambivalent.¹⁷³ Framing the AI Act as an internal market instrument places its legitimacy on the logic of product safety and market integration, not on the direct protection of fundamental rights. This contrasts with other potential legal bases, such as art. 16 TFEU on data protection or art. 19 TFEU on non-discrimination, which were not chosen as primary ones.¹⁷⁴ The result is that the Act's rights-protective provisions are juridically tethered to a market rationale. This may result in a "rule by law" loophole: the use of legal form to mask the absence of substantive protection. There is a danger that the AI Act's legality is of this kind, formally valid, but substantively limited.¹⁷⁵

These limitations are indeed visible also at the national level. The AI Act imposes obligations on providers and deployers but does not require that Member States anchor the use of AI in sensitive sectors in primary legislation.¹⁷⁶ Welfare authorities can deploy fraud-detection algorithms, border agencies can rely on biometric recognition,¹⁷⁷ among others, and courts may experiment with AI,¹⁷⁸ often on the basis of administrative orders or ministerial decrees rather than statutes. From the perspective of the rule

¹⁷¹ As noted also elsewhere, the use of art. 114 TFEU has been considered by scholarship a "catch them all" solution. See O. Pollicino-F. Paolucci, *Digital Constitutionalism*, in L. Floridi-M. Taddeo (eds.), *A Companion to Digital Ethics*, Hoboken, 2025, 313 ff.

¹⁷² See also O. Pollicino-G. Muto, *La legislazione delegata in materia di intelligenza artificiale: la costruzione di una disciplina organica al confine tra scelte governative, controllo parlamentare e vincoli europei*, in *Rivista di Diritto dei Media*, 1, 2025, 393 ff.

¹⁷³ E. Longo, *Grounding media freedom in the EU: the legal basis of the EMFA*, in *Rivista Italiana Di Informatica e Diritto (RIID): Periodico Internazionale Del CNR-IGSG*, 1, 2025, 2 ff.

¹⁷⁴ See the Explanatory Memorandum to the Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts, {Sec(2021) 167 Final} - {Swd(2021) 84 Final} - {Swd(2021) 85 Final}, in part. §2, on the legal basis.

¹⁷⁵ On the difficult interplay between data protection and GenAI see M. Bassini, *DeepSeek, or OpenAI Reloaded? An Italian-Style Approach in Tackling Data Protection in Generative AI*, in *AIRe*, 2(1), 2025, 88 ff.

¹⁷⁶ B. Paola *et al.*, *Le assemblee rappresentative nell'era dell'intelligenza artificiale. Profili costituzionali*, in *Le assemblee rappresentative nell'era dell'intelligenza artificiale. Profili costituzionali*, Milan, 2025, 1 ff.

¹⁷⁷ N. Vavoula, *Regulating AI at Europe's Borders*, in *Verfassungsblog*, 13 December 2024.

¹⁷⁸ See Annex III, AI Act. V. Hendrickx, *Rethinking the judicial duty to state reasons in the age of automation?*, in *Cambridge Forum on AI: Law and Governance*, 1, 2025, 26 ff.

of law, as also conceived by the Venice Commission, this is insufficient: legality demands that such exercises of power have a clear legislative basis, accessible to the public, defining scope, purposes, and safeguards. In its absence, legality risks becoming a purely technical compliance with regulatory checklists, rather than law in the substantive sense.

The Act also strains legality by heavily relying on delegated acts and private standardisation. Art. 97 empowers the Commission to amend Annex III, which lists high-risk AI use cases, through delegated acts.¹⁷⁹ This means that the scope of the Act's strictest obligations can be expanded or narrowed without new legislation. While subject to *ex post* control by Parliament and Council, this delegation nonetheless places significant discretion in the executive. Similarly, the Act relies on harmonised technical standards developed by private standardisation bodies (CEN/CENELEC, ETSI).¹⁸⁰ Providers who comply with these standards are presumed to meet legal requirements. In practice, therefore, the content of obligations such as "appropriate accuracy" or "adequate bias testing" may be determined by technical committees dominated by industry actors. On this view, the AI Act satisfies legality in form but risks undermining it in substance.

6.2. Legal Certainty

With respect to the second benchmark, legal certainty, which demands that laws be clear, foreseeable and stable, enabling individuals to understand and predict legal consequences, in the AI context, this benchmark maps onto issues of scope, transparency, and intelligibility.

The AI Act makes an effort to clarify its scope through a tiered risk model, as mentioned *supra*: prohibited AI practices, high-risk AI, limited-risk AI, and minimal-risk AI. Annex III provides a detailed list of high-risk uses, including biometric identification, recruitment, credit scoring, and access to essential services. This categorisation promotes certainty at the macro level, harmonising what counts as risky across the Union. Yet several uncertainties remain.¹⁸¹ The very definition of "AI system" in art. 3 is broad, encompassing statistical approaches, machine learning, and knowledge-based methods. This breadth is intended to be future-proof,¹⁸² but risks leaving developers unsure whether their systems fall within the Act. More seriously, the scope of "high-risk AI" is not fixed: the Commission can amend Annex III by delegated acts. As a result, providers may face regulatory uncertainty, unsure whether their systems could later be

¹⁷⁹ See also art. 7, AI Act.

¹⁸⁰ M. Cantero Gamito-C. T. Marsden, *Artificial intelligence co-regulation? The role of standards in the EU AI Act*, in *International Journal of Law and Information Technology*, 32(1), 2024, 11 ff.

¹⁸¹ P. Dunn-G. De Gregorio, *The Ambiguous Risk-Based Approach of the Artificial Intelligence Act: Links and Discrepancies with Other Union Strategies*, CEUR Workshop Proceedings, 2022.

¹⁸² European Commission, "Commission Guidelines on the definition of an artificial intelligence system established by Regulation (EU) 2024/1689 (AI Act)", Brussels, 29.7.2025, C(2025) 5053 final.

reclassified as high-risk.¹⁸³ What appears as legal certainty at the level of categories thus masks a dynamic scope that may destabilise predictability. Beyond scope, transparency and explanation are essential to legal certainty. Art. 13 requires providers of high-risk AI to design systems so that users can interpret outputs. Art. 52 mandates informing individuals when they interact with AI. Most significantly, art. 86 establishes a right to explanation for individuals significantly affected by high-risk AI decisions. On paper, this aims to foster certainty: individuals can foresee and understand how decisions are made. But in practice, the provision is limited.¹⁸⁴ It applies only to high-risk AI, only subsidiarily where no other law already grants such a right, and excludes certain domains such as law enforcement biometrics.¹⁸⁵ Moreover, explanations risk being formalistic rather than intelligible. Machine learning systems, especially deep learning models, produce outputs through complex, opaque processes that «rarely correspond to humanly comprehensible reasoning steps».¹⁸⁶ The *Schufa* decision under the GDPR illustrates the problem: while individuals were informed of their credit scores, they could not understand how those scores were calculated, undermining their ability to contest them.¹⁸⁷ The AI Act risks reproducing this dynamic, granting explanations in name but not in substance.¹⁸⁸

Even more striking is the Commission's pervasive reliance on guidelines. While the AI Act entered into application on 2 August 2024, the European Commission has already published guidelines for the interpretation of the AI systems definition under art. 3, on art. 5 on prohibited practices, on the obligations for general-purpose AI models and systemic risk,¹⁸⁹ and guidelines on high-risk obligations are currently under elaboration.¹⁹⁰ These

¹⁸³ For the interplay between high risk and prohibited risk, see European Commission, "Commission Guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (AI Act), Brussels, 29.7.2025 C(2025) 5052 final. See, also, A. Calvi *et al.*, *Contribution to "targeted stakeholder consultation on classification of AI systems as high-risk"*, D.pia.lab, 1 July 2025.

¹⁸⁴ S. Demkova, *The EU's Artificial Intelligence Laboratory and Fundamental Rights*, in M. Fink (ed.), *Redressing fundamental rights violations by the EU*, Cambridge UK, 2025.

¹⁸⁵ Allow to recall F. Paolucci, *Constitutional Safeguards in the Age of AI. A Study on the Fundamental Rights Impact Assessment of Facial Recognition Technology*, Milan, 2025.

¹⁸⁶ M. Garnelo-M. Shanahan, *Reconciling deep learning with symbolic artificial intelligence: representing objects and relations*, in *Current Opinion in Behavioral Sciences*, 29, 2019, 17 ff.

¹⁸⁷ ECJ, C-634/21, cit.

¹⁸⁸ F. Doshi-Velez *et al.*, *Accountability of AI Under the Law: The Role of Explanation*, cit.; E. Falletti, *Credit scoring under scrutiny by the Court of justice of the European union: brief remarks on the «SCHUEFA Decision»*, in SSRN, 14 February 2024, 1 ff.

¹⁸⁹ European Commission, "Commission Guidelines on the definition of an artificial intelligence system established by Regulation (EU) 2024/1689 (AI Act)", Brussels, 29.7.2025, C(2025) 5053 final; European Commission, "Commission Guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (AI Act), Brussels, 29.7.2025 C(2025) 5052 final; European Commission, "Approval of the content of the draft Communication from the Commission – Guidelines on the scope of the obligations for general-purpose AI models established by Regulation (EU) 2024/1689 (AI Act)", Brussels, 18.7.2025 C(2025) 5045 final.

¹⁹⁰ Such guidelines were preceded by a consultation period, as in A. Calvi *et al.*,

documents, though formally non-binding, become *de facto* authoritative: national regulators and courts will look to them for guidance, and providers who deviate from them risk being found non-compliant. This proliferation of soft law generates a form of “moving target” legality: the rules of the game evolve through guidance notes produced by the Commission or the AI Office, outside of the “regular” legislative process.

In this sense, the basics of legal certainty are challenged, since this principle relies on stability and coherence across the legal system. Here, the AI Act introduces potential friction with existing laws, notably the GDPR. For instance, art. 10(5) AI Act allows the processing of special categories of personal data if necessary to detect and correct bias in high-risk AI systems. Yet under the GDPR, processing data under art. 9 requires strict conditions and safeguards.¹⁹¹ This tension creates uncertainty for providers: until clarified by guidance or jurisprudence, actors cannot confidently predict how to comply with both regimes.

In sum, the AI Act promotes certainty at a formal level through harmonisation, but its reliance on broad definitions, delegated amendments, opaque technology, and overlaps with other frameworks means that predictability is often illusory. The law appears clear, but may not be intelligible or foreseeable to those most affected. According to the rule of law’s standards, legal certainty may be formally satisfied in form but not yet in substance.

6.3. Prevention of Abuse of Power

As highlighted in para. 2 of this paper, one of the main purposes of the rule of law is to ensure that (State)¹⁹² power is contained and subject to review, and governance structures designed to avoid arbitrariness. In the AI Act, this benchmark engages both the design of oversight in AI deployments and the concretisation of the enforcement of such rules at the Union level.¹⁹³

Looking at the provisions of the Regulation, art. 14 requires that high-risk AI systems be designed to allow for effective human oversight. This obligation acknowledges the constitutional intuition that legality requires hu-

Contribution to “targeted stakeholder consultation on classification of AI systems as high-risk”, cit., 26.

¹⁹¹ L. Georgieva-C. Kuner, *Article 9 Processing of special categories of personal data*, in C. Kuner *et al.* (eds.), *The EU General Data Protection Regulation (GDPR): A Commentary*, Oxford, 2020.

¹⁹² Nonetheless, the Venice Commission highlights that «the rule of law thus addresses the exercise of power and the relationship between the individual and the state. However, it is important to recognise that during recent years due to globalisation and deregulation there are international and transnational public actors as well as hybrid and private actors with great power over state authorities as well as private citizens», in “Venice Commission Report on the Rule of Law”, 2011, cit., §16.

¹⁹³ A. M. Corrêa *et al.*, *Better together? Human oversight as means to achieve fairness in the European AI Act governance*, in *Cambridge Forum on AI: Law and Governance*, 1, 2025, 29 ff.

man accountability. Yet its content is underspecified.¹⁹⁴ The Act does not define what “effective” oversight entails, nor does it allocate responsibility clearly.¹⁹⁵ In practice, oversight can become ritualistic: a human operator formally “in the loop” but without the information or authority to meaningfully intervene, also by means of lacking the appropriate training.¹⁹⁶ Literature on algorithmic governance has highlighted this risk, describing human oversight as a liability shield rather than a genuine safeguard.¹⁹⁷ From the rule of law perspective, such nominal oversight fails to prevent abuse. To be substantive, «the effectiveness of oversight mechanisms can only be properly assessed and improved when specific goals are clearly articulated in advance»¹⁹⁸ and concretely reviewable.

Even more problematic is the Act’s reliance on codes of practice, particularly for general-purpose AI models.¹⁹⁹ The AI Office is tasked with «the drawing up of codes of practice at Union level in order to contribute to the proper application of this Regulation, taking into account international approaches»,²⁰⁰ inviting also the providers of such models «to participate in the drawing-up of codes of practice».²⁰¹ While co-regulation can bring technical expertise, it also risks regulatory capture.²⁰² Those subject to the

¹⁹⁴ The critics in A. Calvi *et al.*, *Contribution to “targeted stakeholder consultation on classification of AI systems as high-risk”*, cit., 26.

¹⁹⁵ In particular, Recital 73 states that «appropriate human oversight measures should be identified by the provider of the system before its placing on the market or putting into service. In particular, where appropriate, such measures should guarantee that the system is subject to in-built operational constraints that cannot be overridden by the system itself and is responsive to the human operator, and that the natural persons to whom human oversight has been assigned have the necessary competence, training and authority to carry out that role. [...] Considering the significant consequences for persons in the case of an incorrect match by certain biometric identification systems, it is appropriate to provide for an enhanced human oversight requirement for those systems so that no action or decision may be taken by the deployer on the basis of the identification resulting from the system unless this has been separately verified and confirmed by at least two natural persons». It is likely that, as happened also for other obligations, this will be clarified further by a guideline of the European Commission. As explained above, this ritual might spark clarity, in some cases, but also lack of legal certainty, since guidelines, per their name, are not enforceable.

¹⁹⁶ For a taxonomy critique to the rule of law see J. Laux, *Institutionalised distrust and human oversight of artificial intelligence: towards a democratic design of AI governance under the European Union AI Act*, in *AI & Society*, 39(6), 2024, 2853 ff.

¹⁹⁷ M. Fink, *Human Oversight under Article 14 of the EU AI Act*, SSRN Scholarly Paper, Social Science Research Network, Rochester, NY, 2025.

¹⁹⁸ *Ibid.*

¹⁹⁹ The Code of Practice is comprised of three chapters on transparency, copyright and safety and security, and complements the guidelines of the European Commission on the obligations for providers of GPAI models. See European Commission, “Approval of the content of the draft Communication from the Commission – Guidelines on the scope of the obligations for general-purpose AI models established by Regulation (EU) 2024/1689 (AI Act)”, Brussels, 18.7.2025 C(2025) 5045 final.

²⁰⁰ See AI Act, art. 56, para. 1.

²⁰¹ *Ibid.*, para. 2.

²⁰² As noted elsewhere, «today we are witnessing a kind of “code reloaded”, in light of the metamorphosis referred to above, whereby digital intermediaries have transformed from economic operators into full-blown private powers, in competition with public

law participate in shaping its substantive requirements, diluting public accountability. The EU's approach to using codes of conduct as a tool for co-regulation, particularly in the digital domain, reflects an effort to bridge gaps in regulatory frameworks where binding legislation may not yet exist or where enforcement proves challenging. This strategy is especially evident in the '2018 Code of Conduct' aimed at combating disinformation, which initially relied on self-regulation. However, while the initial version was criticised for vague obligations and the lack of measurable indicators to assess platform accountability in content moderation efforts, the revised Code introduced in 2022 incorporated more structured guidelines. This shift represents a movement towards co-regulation, where private entities are encouraged to adhere to certain standards while engaging in dialogue with the European Commission. In this sense, the Digital Services Act (DSA) further institutionalises this approach by recognising codes of conduct as essential tools for co-regulation, with the potential for enforcement consequences if platforms fail to comply with these voluntary commitments. Hence, these experiences demonstrate how codes evolve into hybrid norms, less than law, but more than mere self-regulation, creating a grey zone where accountability is diffuse and democratic oversight is thin. These risks are amplified in the AI context. General-purpose AI providers are not passive intermediaries; they are sophisticated actors capable of shaping information environments and altering the parameters of decision-making across multiple domains. They are, as disinformation research shows, «much less neutral, much less passive and yet much more capable»²⁰³ of influencing content and behaviour, even if they do not resemble traditional publishers.²⁰⁴ Allowing such actors to help define the scope and detail of their obligations under the AI Act blurs the line between regulator and regulated. Particularly, it places constitutional safeguards in the hands of those whose interests may be directly adverse to rights protection.

In this sense, the rule of law requires that powers be exercised under controls external to those being regulated. Codes of practice that invite industry to co-write compliance duties verge on self-regulation, a model that is constitutionally suspect in a field with profound implications for fundamental rights. The AI Act's institutional design compounds these concerns. Alongside codes, it creates a European AI Office within the Commission and a European AI Board composed of national authori-

authorities. Indeed, these operators are distant relatives of the operators that were (not) regulated under European law during the era of digital liberalism». O. Pollicino, *The quadrangular shape of the geometry of digital power(s) and the move towards a procedural digital constitutionalism*, in *European Law Journal*, 29, 1–2, 2023.

²⁰³ O. Pollicino, *The quadrangular shape of the geometry of digital power(s) and the move towards a procedural digital constitutionalism*, *ibid.*

²⁰⁴ As noted by Bassini, «extending constitutional coverage to generative AI content does not prevent state authorities from prohibiting certain types of speech that harm other legal interests (such as individuals' privacy or reputation), but such restrictions would not differ from any other limits on free speech that are constitutionally permissible in the offline world (and also online, via other media)». See M. Bassini, *Speech without a Speaker: Constitutional Coverage for Generative AI Output?*, in *European Constitutional Law Review*, 2025, 1 ff.

ties, both under Commission supervision.²⁰⁵ While such centralisation may bring coordination, it also concentrates power in the executive. Without strong safeguards, the risk is that a framework intended to prevent abuse of AI power will instead entrench new concentrations of institutional and private power with inadequate checks.

6.4. Equality and Non-Discrimination

The fourth benchmark is equality before the law and non-discrimination. AI systems pose acute risks in this domain, and it is, maybe, the most visible of the five rule of law targets: training data may encode structural biases, machine learning models may replicate them, and outputs may disproportionately burden vulnerable groups.²⁰⁶

The AI Act acknowledges these risks by embedding non-discrimination and fairness into its high-risk obligations. Art. 10 requires that training data be relevant, representative, and as free from error and bias as possible: in particular, providers must implement measures to detect and correct biases.²⁰⁷ Art. 27 introduces the Fundamental Rights Impact Assessment (FRIA),²⁰⁸ obliging deployers of certain high-risk AI (notably in public services, credit and insurance) to evaluate impacts across the Charter, including equality and non-discrimination. This is an unprecedented innovation: for the first time, the EU requires *ex ante* assessment of all Charter rights in connection with a technology.

Thus, the FRIA is mandatory for deployers of some high-risk AI systems.²⁰⁹ This obligation extends to deployers governed by public law, private operators providing public services, and operators deploying high-risk systems for the purpose of credit scoring, creditworthiness or risk assessment and pricing of life and health insurance of natural persons.²¹⁰ By mandating FRIAs for these actors, the AI Act emphasises the importance of anticipating and mitigating risks to fundamental rights before deploying high-risk AI systems. More specifically, the FRIA must be performed prior to the initial deployment of the high-risk AI system and should be continued during the whole AI life cycle.²¹¹ This entails a thorough evaluation of the

²⁰⁵ C. Novelli *et al.*, *A Robust Governance for the Ai Act: Ai Office, Ai Board, Scientific Panel, and National Authorities*, in *European Journal of Risk Regulation*, f4, 2024, 1 ff.; F. Paolucci, *The enforcement of the Artificial Intelligence Act: Looking Forward to a Commission Implementing Decisions for Protecting Fundamental Rights*, in O. Pollicino *et al.* (eds.), *La disciplina dell'intelligenza artificiale*, Milan, 2025.

²⁰⁶ R. Xenidis, *Beyond bias: algorithmic machines, discrimination law and the analogy trap*, in *Transnational Legal Theory*, 14(4), 2023, 378 ff.

²⁰⁷ As mentioned above, also using special categories of personal data, as per art. 10, para. 5, AI Act.

²⁰⁸ A. Cosentini *et al.*, *Assessing the Impact of Artificial Intelligence Systems on Fundamental Rights*, cit.

²⁰⁹ AI Act, art. 43, Recital 96.

²¹⁰ AI Act, Annex III, art. 5, lit. b) and c).

²¹¹ AI Act, Recital 96, «Whilst risks related to AI systems can result from the way such systems are designed, risks can as well stem from how such AI systems are used. Deployers

system's potential to adversely affect fundamental rights. For this reason, deployers may rely on previously conducted assessments, including those provided by the system's developer, provided the system's context of use remains unchanged.²¹² However, should any significant modifications arise during the system's use, such as changes in its purpose, affected populations, or identified risks, the deployer is obligated to update the assessment to reflect the new circumstances.²¹³ The AI Act provides limited details on how the FRIA should be conducted, specifying the inclusion of usage descriptions, affected groups of people, and the identification of more vulnerable individuals, as well as potential risks and risk mitigation steps.²¹⁴

The assessment must comprehensively address the following aspects:

System use and purpose: a detailed description of how the AI system will be integrated into the deployer's processes, ensuring alignment with its intended purpose.

Operational parameters: an outline of the timeframe and frequency of the system's deployment.

Affected individuals and groups: an identification of the categories of natural persons and communities likely to be impacted by the system, particularly marginalised or vulnerable groups.

Risk analysis: a detailed evaluation of specific risks of harm that may arise from the system's use, with reference to information provided by the developer in compliance with art. 13 of the AI Act.

Oversight and governance: a description of the human oversight measures that will be implemented, following the developer's guidelines.

Mitigation measures: strategies to address the materialisation of identified risks, including internal governance mechanisms and accessible complaint procedures.

The assessment should then be reported to the market surveillance authority.²¹⁵ A standardised template from the AI Office is under elaboration to ensure consistent evaluations.²¹⁶ From the listed essential requirements, a potential information asymmetry emerges between deployers and providers. Much of the information, such as operational parameters and instructions for use, must be shared by providers to conduct a proper assessment. It is, therefore, needed that the provider and the deployer cooperate, and it is crucial to point out the possible occurrence of contractual inequality between the two actors.²¹⁷ The success of these assessments also relies on

of high-risk AI system therefore play a critical role in ensuring that fundamental rights are protected, complementing the obligations of the provider when developing the AI system. Deployers are best placed to understand how the high-risk AI system will be used concretely and can therefore identify potential significant risks that were not foreseen in the development phase, due to a more precise knowledge of the context of use, the persons or groups of persons likely to be affected, including vulnerable groups».

²¹² AI Act, art. 27, para. 2.

²¹³ *Ibid.*

²¹⁴ AI Act, art. 27, para. 1, lett. a-f).

²¹⁵ AI Act, art. 27, para. 3.

²¹⁶ AI Act, art. 27, para. 5.

²¹⁷ P.G. Chiara-F. Galli, *Normative Considerations on Impact Assessments in EU Digital Policy*, in *Rivista di Diritto dei Media*, 1, 2024.

the transparency and reliability of exchanged information, with providers supplying accurate and comprehensive data.²¹⁸

Yet the substance is weaker than the form. The FRIA is a self-assessment, and not even subject to penalties:²¹⁹ deployers carry it out themselves, with no requirement for independent verification prior to deployment. Results must be communicated to market surveillance authorities, but only after the fact.²²⁰ The risk is that the FRIA becomes defensive paperwork, aimed at minimising liability rather than surfacing risks: without external audit or public scrutiny, the safeguard risks formalism.

Moreover, fairness metrics are inherently contested. Scholars note that different definitions of fairness (equal accuracy across groups, proportionality, lack of disparate impact) can yield divergent outcomes.²²¹ Providers may choose the metric most convenient to them. The Act does not specify which definitions to use, leaving discretion to private actors and standards bodies.²²² In practice, a system could be certified as compliant despite generating significant disparities in outcomes. Equality, as also happens for other “key” words like transparency and fairness, is thereby proclaimed but not enforced. In particular, while the Regulation does prohibit certain AI uses that offend equality outright – e.g. social scoring by public authorities is banned – yet such prohibitions are limited in scope and do not address subtler systemic inequalities.²²³

From the rule-of-law perspective, therefore, the AI Act’s equality provisions risk the same gap as legality and certainty: they satisfy the form of protection without ensuring its substance. The FRIA, in particular, risks devolving into a formalistic exercise unless coupled with independent oversight.

²¹⁸ *Ibid.*

²¹⁹ Notably, the AI Act does not establish a specific penalty regime for violations of art. 27, which requires providers of high-risk AI systems to carry out a FRIA. Under art. 99 AI Act, penalties are defined for breaches of data quality, risk management, or prohibited AI practices, but no express sanction is foreseen for non-compliance with art. 27 obligations. This omission is a potential weakness of the enforcement framework, since the effectiveness of the FRIA depends on the existence of credible deterrents, otherwise is just a void obligation, besides the submission of a report to the AI Office, according to art. 27, para. 3.

²²⁰ P. Chiara-F. Galli, *Normative Considerations on Impact Assessments in EU Digital Policy*, *ibid.*

²²¹ S. Greenstein, *Preserving the rule of law in the era of artificial intelligence (AI)*, in *Artificial Intelligence and Law*, 3- 3, 2022, 291 ff.

²²² A. Kouroutakis, *Rule of law in the AI era: addressing accountability, and the digital divide*, in *Discover Artificial Intelligence*, 4(1), 2024, 115 ff.; M. Rotenberg, *Human Rights Alignment: The Challenge Ahead for AI Lawmakers*, in H. Werthner et al. (eds.), *Introduction to Digital Humanism: A Textbook*, Cham, 2024, 611 ff.; R. Tarchi-A. Gatti, *Intelligenza artificiale e protezione dei dati personali: problemi di metodo e di procedura*, in *DPCE Online*, 64(2), 2024.

²²³ In this sense, the consultation of affected groups is vaguely scatched in the AI Act, as argued in O. Pollicino et al., *AI e sostenibilità Verso una compliance integrata. La valutazione di impatto sui diritti umani tra AI Act e direttiva CSDD*, Milan, 2025.

6.5. Access to Justice

The final benchmark is access to justice, understood as the availability of effective remedies before independent courts. This is the linchpin of the rule of law: without remedies, the other guarantees remain aspirational.²²⁴ The AI Act incorporates remedies in three key provisions:²²⁵ art. 85 grants individuals the right to complain to market surveillance authorities if they suspect an infringement; art. 86 establishes a right to explanation when high-risk AI systems significantly affect an individual; finally, art. 99 para. 10 provides, as a closing provision, that penalties must be subject to judicial remedies and due process. These provisions, taken together, mark an effort to ensure that AI governance is tethered to rights of redress. On paper, they demonstrate that the Union recognises remedies as essential. However, if analysed more thoroughly, their substantive adequacy is dubious.²²⁶ First of all, complaints under art. 85 are not judicial remedies. They are administrative petitions that may trigger regulatory investigations, but the authority has broad discretion whether to act. Even if the complaint is upheld, the individual complainant does not automatically receive redress or reversal of the AI-driven decision that harmed them. Moreover, art. 86, the right to explanation, is the most heralded innovation. It responds to longstanding debates over GDPR art. 22 and the ambiguity surrounding a “right to explanation” of automated decisions.²²⁷ The AI Act makes this explicit: individuals significantly affected by high-risk AI outputs are entitled to «clear and meaningful explanations of the role of the AI system in the decision-making procedure and of the main elements of the decision taken».²²⁸ This, in principle, aligns with rule-of-law demands: decisions must be reasoned to be contestable. But in practice, the provision is narrow in scope and potentially hollow. It applies only to high-risk AI, and only where no equivalent right already exists. Most significantly, «clear and meaningful» is left undefined. In the context of deep learning models, where outputs result from opaque statistical correlations, the risk is that explanations reduce to formal statements (“the AI was used; you did not meet the threshold”) rather than substantive reasons. Such statements do not equip individuals to contest outcomes. This explanation does not suf-

²²⁴ M. Cartabia, *The Rule of Law and The Role of Courts*, in *Italian Journal of Public Law*, 10(1), 2018.

²²⁵ See also Recital 170, « Union and national law already provide effective remedies to natural and legal persons whose rights and freedoms are adversely affected by the use of AI systems. Without prejudice to those remedies, any natural or legal person that has grounds to consider that there has been an infringement of this Regulation should be entitled to lodge a complaint to the relevant market surveillance authority».

²²⁶ M. Eliantonio, *Composite Procedures, the Violation of Fundamental Rights, and the Availability of Sufficient Remedies in the Multi-level EU Judicial Architecture*, in M. Fink (ed.), *Redressing Fundamental Rights Violations by the EU: The Promise of the ‘Complete System of Remedies’*, Cambridge UK, 2024, 345 ff.; P. de Hert, *En Sanctioning Powers and Data Protection: New Tools for Ensuring the Effectiveness of the Gdpr in the Spirit of Cooperative Federalism*, in *EU Law Enforcement*, Oxon, 2021.

²²⁷ S. Demková, *The EU’s Artificial Intelligence Laboratory and Fundamental Rights*, *ibid.*, p 26.

²²⁸ AI Act, art. 86.

fice with the access to justice requirement, which considers not only the mere existence of remedies, but also that they are capable of enabling judicial review. The right to explanation, as discussed, is narrow in scope and potentially unintelligible in practice: while providers must log and document systems, there is no obligation to disclose logs, datasets, or model versions to affected individuals or courts.²²⁹ Moreover, providers routinely invoke trade secrets to shield such information²³⁰.

This problem is not theoretical. In recent cases, it was highlighted by the ECJ how it is complicated for individuals to understand or challenge their credit scores because the algorithmic logic was opaque.²³¹ Finally, art. 99, para. 10, deals with remedies in enforcement, guaranteeing that providers fined under the Act enjoy due process. This is important for ensuring fairness in sanctions, but it protects the regulated entities, not the individuals whose rights are at stake.

Scholars converge on this point. Namely, Nemitz insists that AI must be used according to the law, or not at all.²³² For them, if AI systems cannot be rendered transparent enough to permit review, they should not be deployed. Keller and Drake argue that algorithmic accountability requires proactive contestation: i.e., remedies must be built into design through visibility, explainability, and contestability. The AI Act nods to these principles but does not embed them robustly. The right to explanation is vague; disclosure obligations are absent; contestability remains more aspirational than real. Measured against the Checklist, therefore, the AI Act satisfies access to justice only in form, since, as explained, remedies are proclaimed but deprived of the material conditions that make them real. This deficiency reverberates across the other benchmarks: legality is undermined if courts cannot verify that AI use rests on a statutory basis; legal certainty collapses if explanations are opaque; prevention of abuse of power is hollow if oversight reduces to rubber-stamping, and courts cannot penetrate trade secret shields; equality is aspirational if discriminatory impacts cannot be proven in litigation. The conclusion we can draw from this comparison is that the current form of the AI regulation under the EU's domain sets some principles in form, but without remedy-capable safeguards, it risks hollowing out the rule of law into *rule by law*.

Taken together, the element-by-element analysis reveals a consistent pattern. On each benchmark of the Venice Commission's Checklist, the AI Act offers formal compliance but leaves substantive guarantees fragile: legality anchored in a market basis, certainty displaced into shifting guidance, oversight reduced to ritual, equality left to self-assessment, and remedies proclaimed but deprived of evidence. What emerges is constitutionalisation in form rather than in substance. If the rule of law is to remain

²²⁹ S. Demková, *The EU's Artificial Intelligence Laboratory and Fundamental Rights*, cit., 26.

²³⁰ G. Malgieri-G. Comandé, *Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation*, in *International Data Privacy Law*, 7(4), 2017, 243 ff.

²³¹ ECJ, C-203/22, *CK v Dun & Bradstreet Austria GmbH and Magistrat der Stadt Wien*, 27 February 2025.

²³² P. Nemitz, *Constitutional democracy and technology in the age of artificial intelligence*, in *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 376, 2018, 2018 ff.

a means of rights protection in the digital age, the framework must be refitted to deliver remedies and external controls at the points where AI displaces them. The final section, therefore, turns from diagnosis to prescription, outlining the reforms *de iure condendo* that could align the Act with the Union's constitutional promise.

7. *De Iure Condendo*

This paper has argued throughout that the rule of law in the European Union is not a decorative value but a structural principle whose function is instrumental: it is the means by which rights are guaranteed, arbitrariness is constrained, and power is rendered accountable. As Bin has aptly put it, «the rule of law requires that laws, not men, govern us; the separation of powers and the principle of legality must guarantee this».²³³ When the Union undertakes the constitutionalisation of new technologies, as it has done through the AI Act, the decisive question is not whether rules exist on paper, but whether they are capable of delivering rights protection in practice.

The analysis has shown that when measured against the five benchmarks of the Venice Commission's Rule of Law Checklist, legality, legal certainty, prevention of abuse of power, equality and non-discrimination, and access to justice, the AI Act succeeds in form but fails in substance. Legality is primarily anchored in art. 114 TFEU but detached from rights-oriented Treaty bases. Legal certainty is proclaimed through harmonised categories but undermined by opacity and a proliferation of Commission guidelines. Oversight mechanisms exist but are vague and risk becoming symbolic, while discretion shifts to delegated acts and codes of practice that blur the line between regulator and regulated. Equality is recognised through the FRIA and prohibitions on certain practices, but left to self-assessment without independent verification. Access to justice is formally acknowledged but hollowed out by the absence of intelligible explanations and enforceable disclosure of evidence.

The core finding is that the AI Act constitutionalises AI only in form. It risks hollowing out the rule of law into rule by law: a framework of formal obligations that masks the absence of substantive guarantees. This is not an abstract concern, since automated systems are already determining who receives welfare, who gains access to credit, and who is flagged in policing or migration control. If the safeguards that govern these systems are only nominal, the rule of law itself is trivialised.

Yet this diagnosis also points to the path forward. If the rule of law is to function as a means of rights protection in the age of AI, the framework must be reoriented towards what can be called remedy-capable governance. The Venice Commission's Checklist is particularly useful here because it translates the rule of law into operational benchmarks. By considering them on a policy and future outlook, we can formulate concrete proposals to strengthen the EU's AI governance.

²³³ R. Bin, *Lo Stato di diritto*, cit.

On legality, the Union should move beyond the reliance on art. 114 TFEU as the legal basis, and, to some extent, on art. 16 TFEU, data protection. While harmonisation of the internal market is a legitimate aim, the regulation of technologies that shape fundamental rights should also be anchored in rights-oriented Treaty provisions, such as art. 19 TFEU on non-discrimination. Doing so would clarify that fundamental rights protection is not an ancillary effect of market harmonisation but a core objective of the legislation. At the national level, the Union should require that Member States adopt primary legislation to authorise the deployment of AI in sensitive sectors such as welfare, policing, migration and justice.

On legal certainty, the most urgent reform concerns the reliance on Commission guidelines. The proliferation of soft-law documents, while useful for flexibility, undermines foreseeability. If the meaning of “AI system,” the scope of prohibited practices, or the definition of “meaningful explanation” depends on guidelines that can be updated unilaterally by the Commission or the AI Office, then the law ceases to be stable and predictable. *De iure condendo*, the Union should confine guidelines to clarifications of technical detail, while ensuring that significant interpretative questions are settled in binding legal acts adopted through accountable procedures. In addition, the scope of art. 86 should be broadened beyond Annex III high-risk systems to encompass all algorithmic decisions with significant effects, thereby ensuring that the right to reasons is not contingent on a regulatory classification. In this sense, legal certainty requires not only clarity of categories but intelligibility of decisions, and this can be achieved only by extending and strengthening explanation rights.

On the prevention of abuse of power, reforms should constrain executive discretion and reduce the risks of regulatory capture. The power to amend Annex III through delegated acts should be narrowly circumscribed, with a preference for legislative revision of the high-risk list where possible. Codes of practice, particularly for general-purpose AI, should not be left to providers to co-draft with minimal oversight. Instead, they should be subject to approval and monitoring by independent authorities, with robust participation by civil society and fundamental rights bodies. As a matter of fact, co-regulation must not become self-regulation; the line must be drawn clearly, with constitutional safeguards external to those being regulated. At the same time, art. 14 on human oversight must be clarified: “effective oversight” should be defined as oversight by officials with legal authority to intervene, whose role is documented and reviewable. Only then can oversight be more than ritual.

On equality, the FRIA is a promising tool but needs reinforcement. *De iure condendo*, the FRIA should be subject to independent verification before deployment, at least in the most sensitive sectors. Self-assessment is inadequate where fundamental rights are at stake. Moreover, the law should establish clear and non-negotiable fairness metrics, developed transparently with input from equality bodies and rights experts, rather than leaving providers to select the most convenient definition. Regular external audits should be mandated to ensure that discriminatory impacts are identified and corrected. In addition, the AI Act should be aligned more closely with equality law: where disparate impacts are shown, the burden of proof

should shift to the deployer to demonstrate that the system is justified and proportionate. Without these reforms, equality will remain aspirational. On access to justice, the reforms must be most radical. Remedies must be transformed from formal declarations into effective instruments. This means three things. First, explanation rights must be broadened and deepened. They should apply to all significant AI-driven decisions, and they must provide intelligible reasons, not generic statements. Second, disclosure of evidence must be guaranteed. Providers should be legally obliged to preserve logs, datasets and model versions, and to disclose them to courts and regulators when necessary to review legality, with appropriate safeguards for trade secrets but not at the expense of rights. Third, remedies must be strengthened. The FRIA should not only be a self-assessment but a mechanism subject to independent audit with the power to halt deployment if risks are too great. Together, these reforms would make access to justice substantive. Courts would have the tools to review decisions; individuals would have the information to contest them; oversight bodies would have the authority to intervene.

These proposals are not utopian. They build on existing mechanisms in EU law: the integration of rights-based legal bases, the move from soft law to binding acts, as occurred within the Digital Services Act domain, the experience of independent data protection authorities under the GDPR, and the burden-shifting rules in discrimination law. What is needed is the political will to align AI governance with the Union's own constitutional grammar.

Among the other sources hereby analysed, the 2025 Rule of Law Report shows that the Union already possesses an evaluative grammar for the rule of law, structured around verifiable conditions. Extending this logic to AI governance would mean subjecting the AI Act to a similar cycle of monitoring, using benchmarks of legality, certainty, equality and remedies as indicators of constitutional adequacy. In other words, what the Commission does annually for the Member States should also be done for systemic technologies: constitutionalisation cannot be credible unless the rule of law is made measurable and remedy-capable in practice.

In a nutshell, the broader lesson is that artificial intelligence is the Union's most immediate test of the rule of law. It is not merely another technological field but a transformation in decision-making itself. AI determines how rights are allocated and obligations imposed. It mediates access to welfare, employment, credit, education, migration and justice. If the rule of law is hollow in this domain, it will be hollow in practice. If legality is reduced to delegated acts, certainty to shifting guidelines, oversight to rubber-stamping, equality to self-assessment and remedies to formal complaints, then the rule of law ceases to be a means of rights protection and becomes an empty ritual.

Conversely, if the Union strengthens its framework along the lines suggested, it can demonstrate that the rule of law is resilient, adaptable and substantive. It can show that new forms of power, even when embedded in opaque technical systems, are still subject to law, accountable institutions and effective remedies. In doing so, the Union would not only secure its own constitutional integrity but also shape global governance. The AI

Act is already the first of its kind worldwide. If reoriented towards remedy-capable safeguards, it could become a model for others, setting standards for how AI can be governed in a manner consistent with democracy and rights.

The conclusion, then, is both cautionary and forward-looking. The AI Act is an important step, but it is not sufficient. It constitutionalises AI in form but leaves its substance fragile. Unless reinforced with binding legal bases, clear and stable obligations, independent oversight, verified equality safeguards and effective remedies, it risks trivialising the rule of law into rule by law. Yet the path to reform is clear, and the tools are already within the EU's legal and institutional repertoire. The choice is between a kind of constitutionalisation that is symbolic and another that is real. In the age of artificial intelligence, only the latter will preserve the claim to be a community founded on the rule of law.

Abstract

This article examines the European Union's attempt to regulate artificial intelligence through the AI Act against the Union's own constitutional standard of the rule of law. It advances a functional understanding of the rule of law: not as a symbolic goal, but as a means of protecting rights by providing legality, foreseeability, control of discretion, equality and access to justice. Artificial intelligence poses a structural challenge to this grammar, as autonomy and opacity relocate decision-making from promulgated norms to model design, data selection and inference processes that are difficult to scrutinise. Using the Venice Commission's Rule of Law Checklist as an operational benchmark, the article evaluates the AI Act element by element. It finds that legality is anchored in an internal market basis rather than in rights-oriented Treaty provisions; legal certainty is displaced into shifting Commission guidelines; prevention of abuse is compromised by reliance on delegated acts and industry co-regulation; equality is left to self-assessment through the Fundamental Rights Impact Assessment; and access to justice is nominal, with explanation rights narrow and evidence disclosure absent. The cumulative effect is constitutionalisation in form but not in substance, risking a transformation of the rule of law into *rule by law*. The article concludes by outlining *de iure condendo* reforms that could align the EU AI architecture with its constitutional commitment to the rule of law as a means of protecting rights.

Keywords

AI governance - AI regulation - rule of law - fundamental rights - Artificial Intelligence

Il lato oscuro dell'intelligenza artificiale: come proteggere i dati personali dei lavoratori?*

Marco Capece

Sommario

1. Transizione digitale e *privacy* lavorativa: verso un nuovo paradigma di tutela dei dati personali. - 2. L'intelligenza artificiale nel contesto regolatorio. Opportunità e rischi per la *privacy* nella prospettiva europea e interna. - 3. Decisioni automatizzate e squilibri informativi. La riaffermazione della supervisione umana. - 4. Il contributo degli organismi collettivi alla protezione rafforzata dei dati personali.

1. Transizione digitale e *privacy* lavorativa: verso un nuovo paradigma di tutela dei dati personali

L'evoluzione tecnologica, caratterizzante la transizione digitale, permea le strutture sociali ed economiche al punto da poter parlare di società algoritmica, dai processi decisionali spesso delegati a *software*¹.

L'economia digitale e l'automazione, diffusesi in maniera capillare, paiono promettere l'atteso affermarsi di equità, efficienza e giustizia sociale rispetto al funzionamento dei mercati e delle strutture sociali². Potrebbe rivelarsi

* L'articolo è stato sottoposto, in conformità al regolamento della Rivista, a referaggio "a doppio cieco".

¹ L. Zappalà, *Informatizzazione dei processi decisionali e diritto del lavoro: algoritmi, poteri datoriali e responsabilità del prestatore nell'era dell'intelligenza artificiale*, in *Biblioteca '20 Maggio'*, 2, 2021, 99 ss.; E. Brynjolfsson - A. McAfee, *La nuova rivoluzione delle macchine*, Milano, 2015; L. Zappalà, *Algoritmo* (voce), in S. Borelli - V. Brino - C. Faleri - L. Lazzeroni - L. Tebano - L. Zappalà (a cura di), *Lavoro e tecnologie. Dizionario del diritto del lavoro che cambia*, Torino, 2022, 17 ss.; A. Cipriani - A. Gremolati - G. Mari (a cura di), *Il lavoro 4.0. La quarta rivoluzione industriale e le trasformazioni delle attività lavorative*, Firenze, 2018; F. Rullani - E. Rullani, *Dentro la rivoluzione digitale*, Torino, 2018.

² P. De Petris, *La discriminazione algoritmica. Presupposti e rimedi*, in M. Biasi (a cura di), *Diritto del lavoro e intelligenza artificiale*, Milano, 2024, 225 ss. Dalla digitalizzazione del mercato del lavoro possono derivare benefici tanto in termini di efficienza, inclusività e pari opportunità tanto sul piano dell'incontro tra domanda e offerta, quanto nell'ottica della personalizzazione della formazione professionale, delle politiche attive, dei servizi per il lavoro e dello sviluppo del capitale professionale o anche sul piano della salute e sicurezza sul lavoro. V. anche G. Gaudio, *Le discriminazioni algoritmiche*, in *Lavoro diritti Europa*, 1, 2024, 1 ss.; A. Alaimo, *Il Regolamento sull'Intelligenza Artificiale: dalla proposta della Commissione al testo approvato dal Parlamento. Ha ancora senso il pensiero pessimistico?*, in

«la più grande rivoluzione dell'informazione in tutta la storia umana»³.

L'impiego delle intelligenze artificiali non è, tuttavia, privo di lati oscuri. In particolare, Non risulta, invero, affatto agevole verificare quante e quali informazioni personali siano acquisite ed elaborate mediante l'utilizzo di sistemi automatizzati.

Sul piano prettamente lavoristico, per i datori, avvalersi di sistemi di tal guisa nella gestione del rapporto di lavoro⁴, sin dalla fase precontrattuale nella quale l'IA funge da supporto investigativo e strumento di profilazione del futuro dipendente, qualora gli vengano richieste specifiche qualità attitudinali e professionali, non senza il rischio di sconfinamenti o derive. Ciò impone un'analisi attenta sull'idoneità dell'attuale disciplina giuridica delle tecnologie in parola, in considerazione dell'impatto che hanno sulla realtà⁵ nonché degli effetti negativi, per le persone che lavorano⁶, scaturenti dalla natura accentuatamente pervasiva degli stessi.

Notoriamente, lo studioso di diritto del lavoro si trova a dover rivisitare il proprio sapere, al fine di tenere il passo delle scattanti nuove tecnologie, e, nell'individuare gli aspetti spinosi, anticipa non di rado il legislatore⁷.

Federalismi.it, 18 ottobre 2023.

³ F. Carinci, *Rivoluzione tecnologica e diritto del lavoro: il rapporto individuale*, in *Diritto del lavoro e delle relazioni industriali*, 26, 1985, 203 ss.; E. Dagnino, *Dalla fisica all'algoritmo: una prospettiva di analisi giuslavoristica*, Bergamo, 2019.

⁴ Sul punto, merita evidenziare come, dalla recente ricerca condotta da Anthropic sul mercato USA, emerge un quadro reale molto più sfumato. I dati del rapporto mostrano, anzi, un impiego mirato dell'intelligenza artificiale. Ne è fatto grande uso per lo sviluppo software e per la correzione di *debug* di sistema, per la scrittura tecnica e creativa. In sintesi, l'intelligenza artificiale appare ancora circoscritta al mondo digitale. È la collaborazione a prevalere sull'automazione: Il 57% delle interazioni con l'AI è orientato all'incremento delle capacità umane: apprendimento, iterazione su risultati, perfezionamento di output già prodotti. Sul piano nazionale, dalle rilevazioni effettuate dall'Istat, tra l'altro, emerge che la percentuale di imprese con almeno dieci addetti che impiegata l'AI è cresciuta dal 5 all'8,2% nell'ultimo anno ma è cresciuta in misura inferiore alla media degli altri Paesi europei, pari al 13,5%. Anche in Italia il settore dell'ICT quello in cui è più diffuso il ricorso a ChatGPT e sistemi simili: il 34,6% delle imprese si affida all'AI, ovvero un terzo del totale. Per approfondimenti ulteriori al riguardo, si v. il report *Imprese e Ict - Anno 2024*, pubblicato su *istat.it*; A. Dini, *Intelligenza artificiale e lavoro, quali sono le attività per cui si usa più spesso*, 24 febbraio 2025, in *wired.it*; *AI, ecco chi rischia davvero il posto di lavoro*, in *agendadigitale.eu*. Da ultimo, detta tendenza sembra subire un'inversione. Difatti, secondo i risultati della ricerca *EY Italy AI Barometer 2025*, consultabile su *ey.com*, l'utilizzo dell'AI sul lavoro in Italia è in crescita, passando dal 12% nel 2024 al 46% nel 2025.

⁵ S. Ciucciovino, *La disciplina nazionale sulla utilizzazione della intelligenza artificiale nel rapporto di lavoro*, in *Lavoro diritti Europa*, 1, 2024, 1 ss.; R. Covelli, *Lavoro e intelligenza artificiale: dalla trasparenza alla conoscibilità*, in *Labour and Law Issues*, 9(1), 2023, 78 ss., spec. 80.

⁶ T. Treu, *La digitalizzazione del lavoro: proposte europee e piste di ricerca*, in *Federalismi.it*, 9, 2022, 190 ss.; M. Novella, *Poteri del datore di lavoro nell'impresa digitale: fenomenologia e limiti*, in *Lavoro e diritto*, 3-4, 2021, 451 ss.; E. Signorini, *Il diritto del lavoro nell'economia digitale*, Torino, 2018.

⁷ V. Maio, *Il diritto del lavoro e le nuove sfide della rivoluzione robotica* in *Argomenti di diritto del lavoro*, 2018, 1414 ss.; P.R. Daugherty - H. Wilson, *Human+Machine. Ripensare il Lavoro nell'età dell'intelligenza artificiale*, Milano, 2018; L. Zoppoli, *Il diritto del lavoro dopo l'avvento dell'intelligenza artificiale: aggiornamento o stravolgimento? Qualche (utile) appunto*, in *Biblioteca '20 Maggio'*, 2, 2024, 60 ss. Si veda, sul punto, anche il risalente studio di G. Vardaro, *Tecnica, tecnologia e ideologia della tecnica nel diritto del lavoro*, in *Politica del diritto*, 1986, 75 ss., spec. 118, che vede con largo anticipo le profonde trasformazioni nelle relazioni tra impresa e

La trasformazione digitale⁸ influenza molteplici aspetti della relazione lavorativa⁹, spaziando dalle diverse tipologie di lavoro alle competenze, alla formazione¹⁰; incidendo sul divario digitale¹¹; su tempi, luoghi e carichi di lavoro; sulle condizioni lavorative, sulla retribuzione¹² e sulle modalità di realizzazione della prestazione¹³, con effetti diretti sulla vita privata del dipendente¹⁴.

Vengono alla luce inedite questioni giuridiche che impongono di ricercare un altrettanto inedito equilibrio tra tradizionali valori costituzionali coinvolti, la libertà di iniziativa economica e la tutela dei dati personali del lavoratore¹⁵. Non senza risvolti etici e sociali¹⁶.

Per un verso, infatti, l'innovazione tecnologica può essere un prezioso strumento per «ridurre i costi e accrescere la produttività»¹⁷, creare nuove forme di organizzazione del lavoro e originali opportunità occupazionali¹⁸,

lavoro che le nuove tecnologie informatiche comportano, in particolare segnalando, tra l'altro, lo svanire della titanica figura dell'imprenditore come dominatore della tecnica e organizzatore esclusivo del processo produttivo.

⁸ A. Donini, *Il lavoro attraverso le piattaforme digitali*, Bologna, 2019, 7.

⁹ L. Zappalà, *Le condizioni di lavoro nelle piattaforme: quale strategia europea di tutela del lavoro digitale?* in *Professionalità studi*, IV(1), 2021, 4 ss.; F. Lunardon - E. Menegatti, *I nuovi confini del lavoro: la trasformazione digitale*, 2024, Bologna; V. Bavaro, *Questioni in diritto su lavoro digitale, tempo e libertà*, in *Lavoro e previdenza oggi*, 69(1), 2018, 35 ss.

¹⁰ N. De Angelis, *La formazione dei lavoratori nell'ambiente virtuale. Rischi e opportunità del metaverso*, in *ambienteditto.it*, 4, 2023, 1 ss.; Id., *Il vaccino e la cura. La formazione dei lavoratori come strumento per il mercato*, in *Diritto del mercato del lavoro*, 3, 2022, 611 ss.; G.R. Simoncini, *L'incidenza della rivoluzione digitale nella formazione dei lavoratori*, in *Il lavoro nella giurisprudenza*, 1, 2018, 39 ss.

¹¹ L. Sartori, *Il divario digitale. Internet e le nuove disuguaglianze sociali*, in *Rassegna italiana di sociologia*, 1, 2007, 187 ss.; P. Bozzao, *Longevità lavorativa e politiche di welfare. Nuove sfide e prospettive*, in *Rivista trimestrale di scienza dell'amministrazione*, 1s, 2022, 1 ss.; S. Avveduto, *La società digitale: genere e digital divide*, in *Sociologia della comunicazione*, 57, 2019, 65 ss.

¹² A. Boscati, *Il decent work nel quadro costituzionale*, in M. Brollo - C. Zoli - P. Lambertucci - M. Biasi (a cura di), *Dal lavoro povero al lavoro dignitoso: Politiche, strumenti, proposte*, Bergamo, 2024, 6 ss.; M. Biasi, *Il decent work e la dimensione virtuale: spunti di riflessione sulla regolazione del lavoro nel Metaverso*, in *Lavoro diritti Europa*, 1, 2023, 1 ss.

¹³ G.E. Comes, *Always on. Tecnostress e benessere nell'ambiente di lavoro*, in C. Saggiomo - C. Scialla - A. Tisci (a cura di), *Diritto alla follia. Itinerari storico-giuridici dell'insania*, Collana I quaderni di Heliopolis, 11, Capua, 2022, 141 ss.

¹⁴ R. Zucaro, *Il diritto alla disconnessione. Nuove modalità di tutela della qualità del tempo di vita nella prospettiva giuslavoristica*, in *Lavoro diritti Europa*, 1, 2022, 1 ss.; A. Preteroti, *Ambiente digitale e benessere: la disconnessione come diritto della personalità e strumento di tutela della dignità umana*, in *Ambienteditto.it*, 3, 2023, 1 ss.; C. Timellini, *La disconnessione bussava alla porta del legislatore*, 2019, in *Variazioni su temi di diritto del lavoro*, 1, 2019, 315 ss.

¹⁵ L. D'Arcangelo, *La tutela del lavoratore nel trattamento dei dati personali*, Roma, 2024.

¹⁶ A. Bertolini, *Intelligenza artificiale e responsabilità civile. Problema, sistemi, funzioni*, Bologna, 2025; T. De Mari Casareto Dal Verme, *Intelligenza artificiale e responsabilità civile. Uno studio sui criteri di imputazione*, Trento, 2024.

¹⁷ G. Pigliararmi, *L'impatto dell'intelligenza artificiale sul sistema della sicurezza sociale: problemi e prospettive*, in *Lavoro diritti Europa*, 3, 2024, 1 ss.

¹⁸ V. Van Roy, D. Vertesy - M. Vivarelli, *Technology and employment: Mass unemployment or job creation? Empirical evidence from European patenting firms*, in *Research Policy*, 47, 2018, 1762.

qualità e sicurezza del lavoro¹⁹, oltre a un più equilibrato *work life balance*²⁰. Per un altro, ad allarmare è lo spettro fosco della disoccupazione tecnologica²¹ che «provoca un'incessante redistribuzione delle risorse, umane e materiali, una redistribuzione che coinvolge non solo le nuove leve di lavoratori, ma anche quelle già attive»²², con l'espansione conseguente di alcune attività, ma anche il declino di altre.

Peraltro, non è solo la facilmente ipotizzabile meccanizzazione del lavoro umano a suscitare timori. A preoccupare è, anche, il risultato prodotto dall'interazione uomo-macchina sui diritti dei lavoratori²³.

Al riguardo, nel rilevare l'impatto dei cc.dd. *black boxes*²⁴ sui diritti legati all'espletamento dell'attività lavorativa, è emerso che esistono sistemi automatizzati i cui meccanismi di funzionamento sono ignoti o comunque non comprensibili per l'utilizzatore e per colui che ne subisce l'*output*, ossia il prestatore²⁵.

Ciò consente al datore di lavoro di acquisire e trattare dei dati personali del

¹⁹ In dottrina, vedasi M. Biasi, *Problema e sistema nella regolazione lavoristica dell'intelligenza artificiale: note preliminari*, in *Federalismi.it*, 30, 2024, 162 ss.; C. Romeo, *Tutela e sicurezza del lavoro nell'era della intelligenza artificiale: profili biogiuridici*, in *Il lavoro nella giurisprudenza*, 4, 2024, 445 ss.; F.V. Ponte, *Regole armonizzate sull'intelligenza artificiale: considerazioni lavoristiche*, in *Il lavoro nella giurisprudenza*, 12, 2024, 1093 ss.; A. Tardiola, *Tre quesiti sul rapporto tra sicurezza del lavoro e AI*, in *Lavoro diritti Europa*, 3, 2024, 1 ss.; M. Grandi, *EU-OSHA "Salute e sicurezza sul lavoro nell'era digitale"*, in *Igiene e sicurezza sul lavoro*, 11, 2022, 1 ss.; A. Di Rosa, *Gli effetti della digitalizzazione su salute e sicurezza sul lavoro e la ridefinizione dei confini del benessere del lavoratore a seguito dell'emergenza pandemica*, in *Lavoro diritti Europa*, 3, 2020, 1 ss.

²⁰ L. Calafà, *La conciliazione vita e lavoro in ambiente digitale*, in *Federalismi.it*, 34, 2022, 193 ss.; S. Scarponi, *"Work life balance" fra diritto UE e diritto interno*, in *Biblioteca '20 Maggio'*, 1, 2021, 219 ss.; D. Garofalo, *Contratti di lavoro, mansioni e misure di conciliazione vita-lavoro nel Jobs Act 2*, Bari, 2015; R. Santucci, *Flexicurity e conciliazione tra tempi di vita e di lavoro*, in *Diritti lavori mercati*, 3, 2007, 581 ss.

²¹ A. Perulli - V. Speciale, *Dieci tesi sul diritto del lavoro*, Bologna, 2022, 143 ss.; D. Susskind, *Un mondo senza lavoro*, Milano, 2022; A. McAfee - E. Brynjolfsson, *The Second Machine Age: Work, Progress, and Prosperity in a time of brilliant technologies*, Norton & Co., New York, 2016; OECD, *Automation and independent work in a digital economy*, in *oecd.org*, 2016; R. Campa, *Technological unemployment. A Brief history of an idea*, in *Orbis Idearum*, 2018, 6(2), 57 ss.

²² P. Sylos Labini, *Nuove tecnologie e disoccupazione*, Bari, 1989, 7.

²³ G. Bronzini, *Intelligenza artificiale, impatto sul lavoro e regolazione europea. L'UE mette il broncio al proprio tempo?*, in *Lavoro diritti Europa*, 4, 2024, 1 ss.; M. D'Onghia, *Remotizzazione del lavoro, relazioni sindacali e tutela della salute dei lavoratori*, in L. Zoppoli (a cura di), *Tutela della salute pubblica e rapporti di lavoro*, Napoli, 2021, 251 ss.

²⁴ F. Pasquale, *The Black Box Society*, Harvard University Press, 2016, 59 ss.; D. Castelvechhi, *Can We Open the Black box od AI?*, in *nature.com*, 2016, 21 ss.

²⁵ F. V. Ponte, *Intelligenza artificiale e lavoro. Organizzazione algoritmica, profili gestionali, effetti sostitutivi*, Torino, 2024, 54; L. Zappalà, *Sistemi di IA ad alto rischio e ruolo del sindacato alla prova del risk-based approach*, in *Labour & Law Issues*, 1, 2024, 69 ss.; S. Ciucciavino, *La disciplina nazionale sulla utilizzazione dell'intelligenza artificiale nel rapporto di lavoro*, in *Lavoro diritti Europa*, 1, 2024, 1 ss.; F. Chietera, *Intelligenza artificiale, algoritmi e prove di resistenza giudiziaria*, *Lavoro diritti Europa*, 1, 2024, 7 ss.; M.T. Carinci - S. Giudici - P. Perri, *Obblighi di informazione e sistemi decisionali e di monitoraggio automatizzati (art. 1-bis "decreto trasparenza")*: quali forme di controllo per i poteri datoriali algoritmici, in *Labor*, 1, 2023, 13 ss.; G. Fioriglio, *Intelligenza artificiale, privacy e rapporto di lavoro: una prospettiva informatico giuridica*, in *Lavoro diritti Europa*, 3, 2022, 4 ss.

lavoratore con modalità per quest'ultimo ignote²⁶, con il conseguente instaurarsi di un'asimmetria conoscitiva di non poco rilievo²⁷ e il prospettarsi del rischio di una notevole compromissione dei diritti fondamentali della persona che lavora²⁸.

La dottrina più avveduta sottolinea anche il mutamento dei poteri datoriali che, mediante il ricorso al digitale, subiscono un'estensione²⁹, divengono diffusi³⁰, date le ragguardevoli capacità di indirizzare la prestazione lavorativa, di elaborare e processare dati, nonché di controllare e valutare il personale dipendente³¹, con più ampi margini sanzionatori rispetto al passato. Come non bastasse, l'asimmetria informativa di cui poco sopra abbraccia anche i presupposti fondanti le decisioni automatizzate, accentuando considerevolmente la condizione di subordinazione del lavoratore³². Questi, difettando il sistema informatizzato di chiarezza, finisce per ignorare i limiti legali e contrattuali gravanti sulla figura datoriale, non potendo efficacemente tutelare la propria posizione.

L'attuale e crescente osmosi tra dati, comunità umana e tecnologia³³ rende necessaria una regolamentazione che abbia di mira l'armonizzazione del profilo umano del lavoro con quello tecnologico³⁴.

2. L'intelligenza artificiale nel contesto regolatorio: i rischi per la *privacy* nella prospettiva europea e interna

La combinazione tecnologia-lavoro si presenta come una sorta di "Giano bifronte": da un lato, l'utilizzo di sistemi decisionali automatizzati e di

²⁶ E. Gagnoli, *Il rapporto di lavoro, le decisioni automatiche e i loro limiti. La regolazione di fenomeni incomprensibili*, in A. D'Aloia (a cura di), *Intelligenza artificiale e diritto. Come regolare un mondo nuovo*, Milano, 2020, 470.

²⁷ C. Faleri, *Management algoritmico e asimmetrie informative di ultima generazione*, in *Federalismi.it*, 3, 2024, 228.

²⁸ A. Bellavista, *Sorveglianza elettronica, protezione dei dati personali e tutela dei lavoratori*, in *Lavoro diritti Europa*, 1, 2023, 1 ss.

²⁹ A. Bellavista - R. Santucci (a cura di), *Tecnologie digitali, poteri datoriali e diritti dei lavoratori*, Torino, 2022.

³⁰ M. Novella, *Impresa*, in M. Novella - P. Tullini (a cura di), *Lavoro digitale*, Torino, 2022, 2.

³¹ M. Peruzzi, *Intelligenza artificiale e lavoro. Uno studio su poteri datoriali e tecniche di tutela*, Torino, 2023, 7; P. Tullini, *La questione del potere nell'impresa. Una retrospettiva lunga mezzo secolo*, in *Lavoro e diritto*, 2021, 442 ss.; L. Tebano, *Lavoro, potere direttivo e trasformazioni organizzative*, Napoli, 2020, 193 ss.; L. Zappalà, *Informatizzazione dei processi decisionali e diritto del lavoro: algoritmi, poteri datoriali e responsabilità del prestatore nell'era dell'intelligenza artificiale*, in *WPCSDLE "Massimo D'Antona".IT*, n. 446/2021.

³² L. Zappalà, *Le condizioni di lavoro nelle piattaforme: quale strategia europea di tutela del lavoro digitale?* in *Professionalità studi*, IV(1), 2021, 28.

³³ L'espressione è di M. Faioli, *Data Analytics, robot intelligenti e regolazione del lavoro*, in *Federalismi.it*, 9, 2022, 147 ss.

³⁴ N. De Angelis, *I chiaroscuri dell'intelligenza artificiale nei luoghi di lavoro. Quali prospettive?*, in *Massimario di giurisprudenza del lavoro*, 4, 2024, 768 ss.

monitoraggio risponde all'interesse datoriale, consentendo un'ottimale gestione dei rapporti di lavoro; dall'altro, la perenne e sottile invasione della sfera giuridica dei lavoratori – sotto forma di monitoraggio continuo della prestazione e di accesso ai dati sensibili, elaborati e processati per finalità di gestione e organizzazione del lavoro – grava sui loro diritti e, tra essi, certamente, su quello alla *privacy*³⁵.

Tra i diritti della personalità, quello alla riservatezza è considerato espressione della tutela della dignità della persona³⁶. È un diritto di recente affermazione³⁷.

Prima dell'entrata in vigore della l. 31 dicembre 1996, n. 675, in Italia, alla tutela dei dati personali non era riservata una disciplina specifica, oggi prevista dal Regolamento (UE) 2016/679 e dal d.lgs. 30 giugno 2003, n. 196 e ss.mm.ii. (cd. *Codice Privacy*).

A protezione della riservatezza dei lavoratori, invece, già nel lontano 1970, furono previste le prime disposizioni in materia di *privacy*, rinvenibili nell'art. 8, l. 20 maggio 1970, n. 300³⁸ e, secondo alcuni, anche nell'art. 4 della stessa legge³⁹.

Invero, il ricorso ai sistemi digitalizzati ha rivoluzionato sia le relazioni industriali che quelle tra datore e prestatore. In questi contesti, il concetto di *privacy* assume un'accezione diversa in relazione alla sua espressione (e tutela) nel cyberspazio.

In origine, era sinonimo di privatezza, cioè “spazio” proprio del lavoratore con *ius excludendi alios*⁴⁰; per lasciare, poi, spazio a una nozione differente e di più estesa portata, quale diritto al controllo sulle proprie informazioni. Come si anticipava, rispetto al concreto estrinsecarsi dei rapporti lavorativi, si è spettatori attenti di un epocale mutamento che vede affidate a sistemi digitalizzati funzioni e decisioni tipicamente datoriali, con conseguenti riflessi sulle tradizionali interrelazioni datore-lavoratore e sul sinallagma contrattuale⁴¹.

³⁵ A. Topo, *Automatic management, reputazione del lavoratore e tutela della riservatezza*, in *Lav. dir.*, 3, 2018, 453 ss.; A. Donini, *Profilazione reputazionale e tutela del lavoratore: la parola al Garante della Privacy*, in *Labour Law & Issues*, 3(1), 2017, 36 ss.

³⁶ P. Perlingieri, *La personalità umana nell'ordinamento giuridico*, in P. Perlingieri (a cura di), *La persona e i suoi diritti – problemi del diritto civile*, Napoli, 2005, 13; G. De Simone, *La dignità del lavoro e della persona che lavora*, in *Diritto del lavoro e delle relazioni industriali*, 2019, 633 ss.; L. Gaeta, *La dignità del lavoratore e i “turbamenti” dell'innovazione*, in *Lavoro e diritto*, 2, 1990, 203 ss.

³⁷ F. Carnelutti, *Diritto alla vita privata – contributo alla teoria della libertà di stampa*, in *Rivista trimestrale di diritto e procedura civile*, 1955, 3; S. Sitzia, *Il diritto alla “privatezza” nel rapporto di lavoro tra fonti comunitarie e nazionali*, Padova, 2013. In giurisprudenza, v. Cass. civ., sez. I, 27 maggio 1975, n. 2129, in *CED*.

³⁸ S. Rodotà, *Prefazione*, in A. Trojsi, *Il diritto del lavoratore alla protezione dei dati personali*, Torino, 2013, XIII; S. Sciarra, *Il divieto di indagini sulle opinioni*, in *Rivista trimestrale di diritto e procedura civile*, 1976, 1076 ss.; Id., *Art. 8*, in G. Giugni (diretto da), *Statuto dei lavoratori. Commentario*, Milano, 1979, 101 ss.; S. Niger, *Le nuove dimensioni della privacy. Dal diritto alla riservatezza alla protezione dei dati personali*, Padova, 2006, 52.

³⁹ Pret. Milano, 12 maggio 1972, in *Orientamenti della giurisprudenza del lavoro*, 1972, 260 ss.

⁴⁰ S. Rodotà, *Intervista su privacy e libertà*, Bari, 2005.

⁴¹ S. Ciucciovino, *Tutti i rischi dell'intelligenza artificiale*, in *Collettiva.it*, in *Bollettino Adapt*, 23 febbraio 2023, 2 ss. A. Pizzoferrato, *La disciplina giuridica delle decisioni automatizzate*

Uno dei pericoli insiti nelle nuove forme di organizzazione del lavoro è legato all'identificabilità della figura datoriale con la piattaforma stessa o, finanche, con l'algoritmo che ne regola il funzionamento⁴², con lo sfumare del tradizionale diritto individuale di protezione contro l'invasione di una sfera privata prima chiaramente delimitabile.

Si configura il diritto della persona all'*habeas data*⁴³, ossia al controllo sui propri dati, in considerazione dell'ormai continua e ripetuta cessione d'informazioni, quale presupposto per prendere parte a sociali e per usufruire della fornitura di beni e servizi⁴⁴.

La nozione di *privacy* assume il significato di diritto all'autodeterminazione informativa, la cui tutela si estende alla protezione dei dati attinenti alla persona, per quanto da essa distinti⁴⁵.

Ne è espressione la Carta dei diritti fondamentali dell'Unione europea che riconosce il diritto alla protezione dei dati personali, all'art. 8, tenendola distinta dal diritto al rispetto della vita privata e familiare, riconosciuto all'art. 7. Quest'ultimo ha una natura statica e negativa, mira ad arginare e/o escludere l'invasione altrui dalla dimensione individuale; diversamente, il primo è declinazione dinamica di un diritto di protezione dei dati personali, regola poteri e doveri per un'effettiva protezione dell'individuo. L'attuale sistema normativo multilivello promuove alti livelli di protezione della salute, della sicurezza e dei diritti fondamentali e, tra l'altro, prevede delle disposizioni *ad hoc* per disciplinare l'impiego delle intelligenze artificiali in specifici settori lavorativi.

Sul piano europeo, spiccano il Regolamento (UE) 2016/679 (GDPR) e il Regolamento (UE) 2024/1689, meglio noto come AI Act - che prevedono importanti obblighi di informazione preventiva relativamente ai sistemi decisionali automatizzati che implicano il rilascio di informazioni sui sistemi logici utilizzati e delineano le conseguenze del trattamento per l'interessato -, nonché la direttiva (UE) 2024/2831 del 23 ottobre 2024, sul miglioramento delle condizioni lavorative nelle piattaforme digitali.

Tra le fonti interne, si segnala il d.lgs. n. 152/1997, come modificato, dapprima dal d. lgs. n. 104/2022 (cd. Decreto trasparenza)⁴⁶ e poi dal d.l. 4

nell'ambito del rapporto di lavoro, Torino, 2022.

⁴² Sul punto, cfr. A. Levi, *Il potere di organizzazione nell'economia digitale e i suoi limiti*, in A. Bellavista - R. Santucci (a cura di), *Tecnologie digitali, poteri datoriali e diritti dei lavoratori*, cit., 11 ss.

⁴³ S. Rodotà, *Il mondo nella rete. Quali i diritti, quali i vincoli*, Bari, 2014, 69.

⁴⁴ S. Rodotà, *Il diritto di avere diritti*, Bari, 2012, 396 ss. Si veda anche D. Lyon, *La società sorvegliata. Tecnologie di controllo della vita quotidiana*, Milano, 2001.

⁴⁵ P. Carluccio - R. Ghersi Finocchi, *Il consenso informato quale diritto fondamentale della persona e principio fondamentale di tutela della salute*, in *Giornale di diritto amministrativo*, 3, 2009, 297 ss.; Id., *Principi fondamentali e disciplina del consenso informato*, in *Giornale di diritto amministrativo*, 11, 2009, 1200 ss.; A. Papa, *La problematica tutela del diritto all'autodeterminazione informativa nella big data society*, in *Consulta on line*, 17 aprile 2020, 1 ss.

⁴⁶ In dottrina, si veda D. Garofalo - M. Tiraboschi, *L'impatto sulla disciplina del rapporto di lavoro del d.lgs. n. 104/2022* (emanato in attuazione della direttiva (UE) 2019/1152), in *Adapt Laboru Studies e-book series*, 96, 3 ss., ove si sottolinea l'intervenuta inidoneità della direttiva 91/533/CEE a garantire tali diritti in ragione dei profondi mutamenti dei mercati del lavoro e nonché dall'emersione di nuove forme di lavoro. Si v., inoltre, G.A. Recchia, *Condizioni di lavoro trasparenti, prevedibili e giustiziabili: quando il diritto di informazione sui sistemi automatizzati*

maggio 2023 n. 48, convertito, con modificazioni, con la l. 3 luglio 2023 n. 85 (cd. Decreto lavoro). L'art. 1 *bis* pone specifica attenzione sulla posizione di debolezza del dipendente e predispone all'uopo particolari forme di protezione e tutela ispirate al principio di trasparenza algoritmica, allargando il perimetro delle informazioni che il datore è tenuto a fornire ai lavoratori. Tra queste figurano le informazioni relative alle modalità di funzionamento dell'algoritmo e alla struttura dei sistemi decisionali, di monitoraggio e di valutazione delle prestazioni delle IA.

Riguardo ai sistemi decisionali integralmente automatizzati – mediante una procedura simile a quella prevista nell'AI Act – è imposto che il datore supervisioni il corretto funzionamento del sistema e la compatibilità dello stesso con i diritti fondamentali del lavoratore⁴⁷, per evitare il verificarsi di un danno ingiusto alla persona coinvolta⁴⁸.

Va, peraltro, evidenziato che, ancora prima dell'entrata in vigore dell'AI Act⁴⁹, il Consiglio dei Ministri aveva approvato il disegno di legge in materia di intelligenza artificiale⁵⁰, che dettaglia il disposto normativo euro-unitario e prevedendo una serie di ambiti di interesse rilevanti interconnessi alle intelligenze artificiali, tra cui figura quello del lavoro.

In particolare, l'art. 11 individua i possibili benefici conseguibili dall'impiego delle intelligenze artificiali, tra cui l'auspicato miglioramento delle condizioni di lavoro, la tutela dell'integrità psico-fisica dei lavoratori, l'accrescimento della qualità delle prestazioni lavorative e della produttività delle persone, conformemente al diritto comunitario.

Tale elencazione corrisponde ad una scelta finalistica che pare vincolarne l'utilizzo alle sole finalità individuate, escludendo ulteriori ambiti di applicazione delle IA nei contesti lavorativi⁵¹, ma ciò finirebbe per collidere con il principio della libera iniziativa economica (art. 41 Cost.; art. 6 TFUE) e con gli stessi principi ispiratori dell'AI Act, volti a porre in equilibrio i diritti fondamentali ed il progresso tecnologico⁵².

In ogni caso, le garanzie di sicurezza, affidabilità, trasparenza impongono il rispetto dei beni giuridici fondamentali della dignità umana, nonché della riservatezza dei dati personali. In quest'ottica, per il datore di lavoro o il committente sarà doveroso informare il lavoratore dell'utilizzo dell'intel-

diventa uno strumento di tutela collettiva, in *Labour & Law Issues*, 1, 2023, 40.

⁴⁷ A. Viscomi, *Per una sandbox giuslavoristica. Brevi note a partire dal "decreto trasparenza"*, in *Labour & Law Issues*, 9(1), 2023, 129 ss.

⁴⁸ E. Dagnino, *Il diritto interno: I sistemi decisionali e di monitoraggio automatizzati tra trasparenza e coinvolgimento*, in M. Biasi, (a cura di) *Diritto del lavoro e intelligenza artificiale*, Milano, 2024, 157 ss.; A. Donini, *Informazione sui sistemi decisionali e di monitoraggio automatizzati: poteri datoriali e assetti organizzativi*, in *Diritto lavori mercati*, 1, 2023, 88 ss.; M. Marazza - F. D'aversa, *Dialoghi sulla fattispecie dei "sistemi decisionali o di monitoraggio automatizzati" nel rapporto di lavoro (a partire dal Decreto trasparenza)*, in *Giustiziacivile.com*, 2022, 8 ss.

⁴⁹ Risalente al primo agosto 2024.

⁵⁰ Nel CDM n. 78 del 23 aprile 2024 il Governo ha approvato il ddl avente ad oggetto "Disposizioni e delega al Governo in materia di intelligenza artificiale".

⁵¹ E. Dagnino, *Intelligenza Artificiale e diritto interno: note a margine del disegno di legge di iniziativa governativa*, *Lavoro diritti Europa*, 4, 2024, 5 ss.

⁵² Cfr. M. Biasi, *Problema e sistema nella regolazione lavoristica dell'intelligenza artificiale note preliminari*, in *Federalismi.it*, 30, 2024, 172.

ligenza artificiale nei casi e con le modalità di cui all'art. 1-*bis* del d. lgs. 26 maggio 1997, n. 152⁵³.

Dette informazioni devono essere rese intellegibili per il lavoratore, non devono richiedere una spiegazione complessa degli algoritmi utilizzati o la loro integrale divulgazione che restano nella sfera della proprietà industriale⁵⁴.

Ai sensi dell'art. 12, è prevista l'istituzione, presso il Ministero del lavoro e delle politiche sociali, dell'Osservatorio sull'adozione di sistemi di intelligenza artificiale nel mondo del lavoro proprio per massimizzarne i benefici e contenerne i rischi, definirne una strategia d'utilizzo, monitorarne l'impatto sul mercato del lavoro e identificare i settori lavorativi che ne sono maggiormente interessati, per promuovere la formazione dei lavoratori e dei datori.

In ambito europeo, l'AI Act contiene un modello normativo volto ad armonizzare le legislazioni per l'attuazione del mercato interno e assicurare un'efficace protezione ai diritti fondamentali della persona nei diversi ambiti di applicazione⁵⁵. Qualifica i sistemi di IA utilizzati nel settore dell'occupazione⁵⁶ ad alto rischio⁵⁷, individuando un programma di valutazione, gestione e controllo *ex ante* dei rischi noti e ragionevolmente prevedibili in condizioni di corretto utilizzo dei sistemi e nelle ipotesi di uso improprio. Il corretto funzionamento del modello è frutto dall'interconnessione tra fornitore e *deployer*⁵⁸ (datore di lavoro), per la trasmissione di informazioni rilevanti sul funzionamento del sistema e dei relativi rischi per la salute e la sicurezza e per i diritti fondamentali.

Di recente, la direttiva (UE) 2024/2831 del 23 ottobre 2024, oltre a specificare e integrare i diritti preesistenti in materia di protezione dei dati personali, incentiva il dialogo sociale sui sistemi basati su algoritmi, prevedendo diritti collettivi in materia di informazione e consultazione⁵⁹ al fine di assicurare equità, responsabilità nella gestione algoritmica e maggiore trasparenza e comprensione dei processi di gestione algoritmica, unitamente a un accesso più efficace agli strumenti per opporsi alle decisioni automatizzate, con conseguente miglioramento delle condizioni di lavoro e del controllo dei propri dati⁶⁰.

⁵³ Art. 11, n. 2.

⁵⁴ M. Bassini, *Intelligenza Artificiale e diritti fondamentali: considerazioni preliminari*, in M. Biasi (a cura di), *Diritto del lavoro e intelligenza artificiale*, Milano, 2024, 23 ss.

⁵⁵ Considerando 20 del Regolamento (UE) 2024/1689.

⁵⁶ Considerando 57 dell'AI Act.

⁵⁷ M. Peruzzi, *La discriminazione algoritmica*, in *Equal - Rivista di diritto antidiscriminatorio*, 1, 2024, 10 ss.; A. Aloisi - V. De Stefano, *Il nuovo Regolamento Ue sull'intelligenza artificiale e i lavoratori*, in *rivistailmulino.it*, 2021.

⁵⁸ M. Peruzzi, *Intelligenza artificiale e lavoro: l'impatto dell'AI Act nella ricostruzione del sistema regolativo Ue di tutela*, in M. Biasi (a cura di), *Diritto del lavoro e intelligenza artificiale*, Torino, 2024, 115 ss.

⁵⁹ M. Delfino, *Lavoro mediante piattaforme digitali, dialogo sociale europeo e partecipazione sindacale*, in *Federalismi.it*, 25, 2023, 175 ss.

⁶⁰ A. Donini, *Tecniche avanzate di analisi dei dati e protezione dei lavoratori*, in *Diritto del lavoro e delle relazioni industriali*, 1, 2018, 222 ss.; C. Spinelli, *La trasparenza delle decisioni algoritmiche nella proposta di direttiva UE sul lavoro tramite piattaforma*, in *Lavoro diritti Europa*, 2, 2022, 1 ss.

È evidente che l'assetto di tutele nazionale e sovranazionale sul trattamento dei dati si ispira alla realizzazione, se non della totale trasparenza, quantomeno della reperibilità dei responsabili del trattamento e del diritto di decidere della sorte dei propri dati personali⁶¹.

Per rendere effettiva la tutela, è indispensabile che la disciplina a protezione dei dati personali si evolva, soprattutto in una fase storica in cui tali dati sono oggetto di commercializzazione⁶².

3. Decisioni automatizzate e squilibri informativi. La riaffermazione della supervisione umana

Su di un piano squisitamente lavoristico, la *gig economy* basa il proprio funzionamento su di un meccanismo d'intermediazione algoritmico – ben distante dalla disciplina in materia di somministrazione e agenzia – è realizzato mediante l'impiego di sistemi d'intelligenza artificiale, diffusamente confusi con la figura datoriale⁶³ e tendenti ad eludere le tipiche garanzie lavoristiche.

Ciò non deve, tuttavia, impedire di ricondurre ai tradizionali canoni lavoristici il nuovo, digitalizzato, atteggiarsi dei poteri datoriali proprio perché, come si è detto⁶⁴, la pervasività e la spersonalizzazione della gestione dei rapporti di lavoro comportano rischi di accentuazione della debolezza contrattuale del lavoratore e possibili nuove forme di discriminazione⁶⁵.

Si pensi, in particolare, ai *bias*, cioè agli errori commessi dalla strumentazione intelligente in fase di apprendimento e capaci di produrre risultati influenzati da dati di addestramento erranei o da pregiudizi umani cui seguono decisioni discriminatorie⁶⁶. Si pensi pure ai casi in cui l'azienda si avvalga, all'insaputa dei lavoratori, di un sistema di rete che memorizza le attività svolte dai dipendenti alle proprie postazioni di lavoro per registrarne le condotte illecite o accedere a informazioni su fatti irrilevanti ai fini della prestazione⁶⁷.

Questi rischi dovrebbero trovare un argine nella comune visione antropocentrica dei rapporti di lavoro dei regolamenti (UE) 2024/1689 (AI Act) e 2016/679, da realizzare innanzitutto ricorrendo all'adozione sistemi di IA

⁶¹ R.Covelli, *Lavoro e intelligenza artificiale: dalla trasparenza alla conoscibilità*, cit., 80.

⁶² S. Zuboff, *Il capitalismo della sorveglianza. Il futuro dell'umanità nell'era dei nuovi poteri*, Roma, 2019.

⁶³ A. Aloisi - V. De Stefano, *Il tuo capo è un algoritmo. Contro il lavoro disumano*, Milano, 2020.

⁶⁴ Cfr., *infra*, par. 1.

⁶⁵ C. Alessi, *Lavoro tramite piattaforma e divieti di discriminazione nell'UE*, in C. Alessi - M. Barbera - L. Guaglianone (a cura di), *Impresa, lavoro e non lavoro nell'economia digitale*, Bari, 2019, 695.

⁶⁶ E. Stradella, *Stereotipi e discriminazioni: dall'intelligenza umana all'intelligenza artificiale*, in *Consulta on line*, 30 marzo 2020, 1 ss.

⁶⁷ Cfr. Provvedimento del Garante della Privacy, n. 10096474 del 12 dicembre 2024 secondo cui ai sensi dell'art. 58 par. 2 lett. f) e 83 del regolamento 2016/679 è vietato al datore di lavoro di trattare dati riferiti ai dipendenti previsti nel "Disciplinare interno per un corretto utilizzo di strumenti informatici, posta elettronica e navigazione in internet".

che non siano lesivi della dignità dei lavoratori⁶⁸, e, in secondo luogo, alla configurazione dell'obbligazione del *deployer* di effettuare una valutazione di impatto per la protezione dei dati (DPIA)⁶⁹, indipendentemente dall'utilizzo di sistemi decisionali o comunque automatizzati⁷⁰.

Si rende doverosa, per il titolare del trattamento, non solo l'adozione di tutte le misure idonee ad assicurare la conformità ai principi del GDPR e la trasparenza, ma soprattutto la valutazione, con la mitigazione e la gestione dei rischi per tutti i diritti e le libertà degli interessati. La DPIA include, altresì, «una valutazione della necessità e della proporzionalità del trattamento in relazione alle finalità» (art. 35, par. 7, lett. b). Si ritiene che il datore debba giustificare la scelta di un certo tipo di algoritmo rispetto a un determinato obiettivo organizzativo, a fronte dell'assenza di alternative meno invasive per i diritti e le libertà coinvolti.

In linea con questa impostazione, la direttiva UE in materia di lavoro su piattaforme digitali⁷¹, prevede, all'art. 10, che gli Stati membri impongano alle piattaforme digitali di valutare costantemente l'impatto delle decisioni prese o sostenute dai sistemi automatizzati sulle persone e sulle loro condizioni di lavoro⁷², per poi individuare e applicare le misure necessarie ad evitare le violazioni dei diritti dei lavoratori.

Dunque, per quanto l'algoritmo possa surrogarsi al datore di lavoro nell'esercizio di alcuni suoi poteri, esso resta «strumento dell'azione umana»⁷³, con funzione di mero supporto.

⁶⁸ A. Sartori, *L'impatto dell'intelligenza artificiale sul controllo e la valutazione della prestazione, e sull'esercizio del potere disciplinare*, in *Lavoro diritti Europa*, 3, 2024, 7 ss.; A. Ingraio, *Il controllo a distanza sui lavoratori e la nuova disciplina privacy: una lettura integrata*, Bari, 2018; nonché Id., *Il potere di controllo a distanza sull'attività lavorativa e la nuova disciplina della privacy nella sfida delle nuove tecnologie*, in G. Ludovico - F. Fita Ortega - T. Nahas (a cura di), *Nuove tecnologie e diritto del lavoro*, Milano, 2021, 111 ss.; R. Messinetti, *La Privacy e il controllo dell'identità algoritmica*, in *Contratto e impresa Europa*, 1, 2021, 121 ss.

⁶⁹ Com'è previsto ai sensi dell'art. 35 del GDPR.

⁷⁰ L'AI Act, all'art. 26, par. 9, specifica che il *deployer* utilizza le informazioni ricevute dal fornitore per adempiere all'obbligo di valutazione d'impatto sulla protezione dei dati personali (DPIA) previsto dal GDPR. Sulla medesima linea di raccordo, l'art. 27, par. 4, stabilisce che se uno qualsiasi degli obblighi ivi previsti è già rispettato a seguito della DPIA, la valutazione d'impatto sui diritti fondamentali integra la valutazione condotta a norma del GDPR.

⁷¹ Direttiva (UE) 2024/2831 del Parlamento Europeo e del Consiglio del 23 ottobre 2024 relativa al miglioramento delle condizioni di lavoro nel lavoro mediante piattaforme digitali. Il tema del lavoro su piattaforma è stato tra quelli che maggiormente hanno affascinato la dottrina. Sul punto, v. E. Menegatti - G. Bandelloni - M. Giovannone - F. Di Noia - M. Dolores Ferrara, in M. Biasi (a cura di), *Diritto del lavoro e intelligenza artificiale*, Milano, 2024, 455 ss.

⁷² Cfr. sul punto M. Peruzzi, *LA e obblighi datoriali di tutela del lavoratore: necessità e declinazioni dell'approccio risk based*, in *Federalismi*, 30, 2024, 236.

⁷³ È della stessa idea S. Ciucciovino, *La disciplina nazionale sulla utilizzazione dell'intelligenza artificiale nel rapporto di lavoro*, cit., 8, secondo la quale non si può configurare la macchina intelligente come un *procurator* o delegato del datore, quest'ultimo, infatti, risulta essere pienamente titolare e responsabile dell'atto decisionale che assume. Diverso è il pensiero di U. Gargiulo, *Intelligenza artificiale e poteri datoriali: limiti normativi e ruolo dell'autonomia collettiva*, in *Federalismi.it*, 28 giugno 2022, secondo cui si assiste a un mutamento morfologico dei poteri e il relativo esercizio appare quasi rarefatto o talvolta assente.

4. Il contributo degli organismi collettivi alla protezione rafforzata dei dati personali.

Tra le strategie normative di bilanciamento dell'innovazione e dei diritti fondamentali, non è trascurabile il ruolo degli organismi collettivi.

Al riguardo, l'AI Act, all'art. 26, par. 7, impone ai *deployer*⁷⁴, (datori di lavoro), di informare i rappresentanti dei lavoratori e i lavoratori interessati che saranno soggetti all'uso del sistema di IA ad alto rischio, prima di mettere in servizio o utilizzare un tale sistema di intelligenza artificiale sul luogo di lavoro⁷⁵.

Questa previsione è espressione di un'opzione regolativa finalizzata a spingere gli Stati membri meno pronti in tema di *employee's involvement* ad aprirsi alla necessità di informare, ma anche di consultare per negoziare⁷⁶. La direttiva (UE) 2024/2831 intensifica l'obbligo di informazione in favore dei rappresentanti dei lavoratori, richiedendo la loro partecipazione al procedimento di sorveglianza umana di cui all'art. 10 e a quello di "valutazione d'impatto delle decisioni assunte o sostenute dai sistemi decisionali o di monitoraggio automatizzati sulle persone che svolgono un lavoro mediante piattaforme digitali".

Del medesimo tenore l'art. 1 *bis* del d.lgs. n. 152/1997 che prevede l'obbligo datoriale di informare gli organismi sindacali aziendali di rappresentanza dei lavoratori e, in assenza di essi, le sedi territoriali delle associazioni sindacali comparativamente più rappresentative sul piano nazionale⁷⁷, per renderle edotte sull'impiego di sistemi integralmente automatizzati.

Prima facie, gli organismi sindacali sembrerebbero rivestire un ruolo secondario⁷⁸, di mera "informazione" sull'utilizzo dei sistemi decisionali e di monitoraggio automatizzati⁷⁹.

⁷⁴ Sui diritti e obblighi del datore di lavoro in qualità di *deployer*, si rinvia a M. Peruzzi, *Intelligenza artificiale e lavoro: l'impatto dell'AI Act nella ricostruzione del sistema regolativo UE di tutela*, cit., 115 ss.

⁷⁵ In tal senso si pone anche la direttiva (UE) 2024/2831, relativa al miglioramento delle condizioni di lavoro nel lavoro mediante piattaforme digitali, che prevede, all'art. 9, misure finalizzate a garantire al prestatore di lavoro informazioni adeguate sull'utilizzazione, da parte della piattaforma, di sistemi di monitoraggio automatizzato sulla sua attività lavorativa come pure sui sistemi di decisione gestiti interamente dall'algoritmo. Per un commento, si veda T. Treu, *CRSD, direttiva sui lavoratori delle piattaforme e valutazione dei rischi*, in *Federalismi.it*, 18 dicembre 2024; M. Biasi, *La presunzione relativa di subordinazione nella "Direttiva Piattaforme": brevi spunti*, in *Lavoro diritti Europa*, 4, 2024, 1 ss.

⁷⁶ Così A. Alaimo, *Il Regolamento sull'Intelligenza Artificiale: dalla proposta della Commissione al testo approvato dal Parlamento. Ha ancora senso il pensiero pessimistico?*, cit., 3; M. Aimò, *Trasparenza algoritmica nel lavoro su piattaforma: quali spazi per i diritti collettivi nella proposta di Direttiva in discussione?*, in *Lavoro diritti Europa*, 2, 2024, 5; M. Biasi, *Problema e sistema nella regolazione lavoristica dell'intelligenza artificiale: note preliminari*, in *Federalismi.it*, 30, 2024, 183.

⁷⁷ Si veda S. Ciucciovino, *La disciplina nazionale sulla utilizzazione dell'intelligenza artificiale nel rapporto di lavoro*, cit., 1 ss.

⁷⁸ M. Corti, *Potere di controllo e nuove tecnologie. Il ruolo dei partner sociali*, in *Labour & Law Issues*, 1, 2023, 70 ss.

⁷⁹ Cfr. A. Alaimo, *Nuovi termini del discorso giuridico sulla partecipazione dei lavoratori. Partecipazione organizzativa e diretta versus partecipazione indiretta: per una dialettica senza contrapposizione*, in *Biblioteca '20 Maggio'*, 1, 2024, 36.

In realtà, è proprio l'autonomia collettiva a dovere assumere un ruolo attivo per sviluppare e declinare i diritti di nuova generazione o reinterpretare principi e criteri di tutela classici del diritto del lavoro, per ridisegnare nuovi equilibri⁸⁰.

Sarebbe, all'uopo, fondamentale attenuare l'asimmetria informativa, accrescere la trasparenza, specialmente attraverso una partecipazione all'analisi e alla valutazione di impatto dell'uso dell'IA nei contesti lavorativi⁸¹.

Ai limiti del richiamato art. 1-*bis* supplisce la giurisprudenza, che considera l'informazione sulle modalità di funzionamento dei processi decisionali automatizzati oggetto di un diritto sia soggettivo che collettivo⁸², giudizialmente tutelabile, e possibile, efficace ausilio nei giudizi antidiscriminatori di nuova generazione.

A rafforzare e riconoscere valore alla partecipazione collettiva, il riconosciuto carattere antisindacale del rifiuto del datore di lavoro di fornire le informazioni riguardanti l'utilizzo di sistemi decisionali e di monitoraggio automatizzati⁸³.

⁸⁰ G. Proia - R. Pessi (a cura di), *Autonomia collettiva e realtà economica*, Torino 2025; S. Muccio - A. Stanchi, *Individui, tecnologia e lavoro: rischi emergenti, tutele attuali e prospettive evolutive nell'era dell'automatizzazione*, in A. Mortara - R. Scramaglia, (a cura di), *Cambiamenti in un mondo instabile. Ambiente, tecnologia e consumi*, Milano, 2023, 288; E. Balletti, *Trasparenza e contrattazione collettiva*, in *Massimario di giurisprudenza del lavoro*, 4, 2023, 650 ss.; L. Lorea, *Protezione del lavoratore nell'impresa tecnologicamente organizzata e ruolo del sindacato*, in *Lavoro diritti Europa*, 3, 2022, 1 ss.

⁸¹ L. Imberti, *Intelligenza artificiale e sindacato. Chi controlla i controllori artificiali?*, cit., 20.

⁸² V. Trib. Palermo, 3 aprile 2023, in *Onelegale* che considera antisindacale il diniego da parte di una società di *food delivery* di fornire alle organizzazioni sindacali informazioni sull'utilizzo di sistemi decisionali o di monitoraggio automatizzati, con conseguente ordine di fornire le informazioni richieste.

⁸³ I provvedimenti più rilevanti in materia sono stati emessi dal Tribunale di Palermo in data 4 aprile 2023 (14491/2023) e in data 20 giugno 2023 in relazione, rispettivamente, ai casi "Uber Eats" e "Foodinho. I giudici palermitani confermano che il diniego del datore di lavoro di fornire l'informativa di cui all'art. 1 *bis* del d. lgs. n. 152/1997 alle rappresentanze dei lavoratori costituisca una condotta antisindacale. Secondo il decreto n. 14191/2023 del 4 aprile 2023 "le informazioni di cui all'art 1 *bis* del d.lgs. n. 152 del 1997 possono essere richieste "anche" dalle associazioni per cui il rifiuto datoriale integra una lesione del diritto alla informativa azionabile anche da parte sindacale, quindi in aggiunta e non in alternativa rispetto all'eventuale previo rilascio al lavoratore". Sul punto, si vedano anche Trib. Torino 12/03/2024, n. 231, in *deJure*; Trib. Bari, 4 marzo 2025, n. 852, in *deJure*. In dottrina, per un'analisi, cfr. E. Lackova, *Opacità degli algoritmi e decreto trasparenza: il sindacato fa la sua parte*, in *Rivista italiana di diritto del lavoro*, 3, 2023, 367 ss.; G.A. Recchia, *Condizioni di lavoro trasparenti*, cit., 17 ss.; M. Marazza - F. D'Aversa, *Dialoghi sulla fattispecie dei "sistemi decisionali o di monitoraggio automatizzati" nel rapporto di lavoro (a partire dal Decreto trasparenza)*, 8 novembre 2022, in *Giustiziavivile.com*; A. Zilli, *La trasparenza nel lavoro subordinato. Principi e tecniche di tutela*, in *Labor. Istantanee*, Pisa, 2022; C. Spinelli, *La trasparenza delle decisioni algoritmiche nella proposta di direttiva UE sul lavoro tramite piattaforma*, in *Lavoro diritti Europa*, 2, 2022, 2 ss.; G. Proia, *Origine, evoluzione e funzioni della trasparenza nei rapporti di lavoro*, in *Massimario di giurisprudenza del lavoro*, 4, 2023, 719 ss.; M.T. Carinci - S. Giudici - P. Perri, *Obblighi di informazione e sistemi decisionali e di monitoraggio automatizzati (art. 1-bis "Decreto Trasparenza"): quali forme di controllo per i poteri datoriali algoritmici?*, in *Labor*, 2023, 7 ss.; G. Peluso, *Obbligo informativo e sistemi integralmente automatizzati*, in *Labour & Law Issues*, vol. 9, 2, 2023, 100 ss.; C. Timellini, *Quale reale trasparenza nel rapporto di lavoro con gli ultimi adempimenti?*, in *Variazioni su temi di diritto del lavoro*, 2, 2023, 571 ss. Con l'entrata in vigore della nuova legge del 23 settembre 2025, n. 132 sull'intelligenza artificiale, peraltro, dal 10 ottobre 2025, è previsto l'obbligo per il datore di lavoro ed il committente di

Appare imprescindibile un sempre maggiore coinvolgimento degli organismi collettivi, in funzione di contrappeso alle diseguaglianze e alle criticità descritte, per maturare un nuovo approccio alle intrinseche potenzialità organizzative dell'IA, sì da coglierne appieno le vantaggiose sfaccettature.

Abstract

L'autore, rilevata la capillare espansione dell'evoluzione tecnologica determinata dalla transizione digitale, pone l'attenzione sull'inedita complessità connesse all'esercizio automatizzato dei poteri datoriali e sui riflessi che esso ha sui diritti fondamentali del lavoratore. Si sofferma sulla disciplina multilivello del trattamento dei dati personali, alla ricerca di un bilanciamento tra la libertà d'iniziativa economica privata e la tutela della *privacy* del prestatore di lavoro. Auspica delle soluzioni che assicurino una protezione rafforzata dei dati personali, rilevando pregi e criticità delle intelligenze artificiali.

The author, noting the widespread expansion of technological evolution determined by the digital transition, focuses on the unprecedented complexity connected to the automated exercise of employer powers and the implications it has on the fundamental rights of the worker. He focuses on the multilevel regulation of the processing of personal data, seeking a balance between the freedom of private economic initiative and the protection of the privacy of the employee. He hopes for solutions that ensure enhanced protection of personal data, noting the advantages and criticalities of artificial intelligence.

Keywords

transizione digitale - privacy - lavoratori - tutela - intelligenza artificiale

La convenzione ONU sul crimine informatico: tra obiettivi globali e tensioni geopolitiche*

Elisabetta Pino

Sommario

1. Premessa - 2. Il tortuoso percorso negoziale verso un consenso globale. - 3. Architettura e profili innovativi della Convenzione ONU. - 4. La Convenzione ONU e la Convenzione di Budapest: due modelli a confronto. - 5. Le perplessità delle organizzazioni internazionali alla ricerca di un equilibrio tra cooperazione internazionale e tutela dei diritti. - 6. Conclusioni.

1. Premessa

L'adozione da parte dell'Assemblea Generale delle Nazioni Unite della Convenzione sul Cybercrime nel dicembre 2024 costituisce un punto di svolta nel quadro istituzionale internazionale, poiché si propone di colmare un vuoto regolativo attraverso la codificazione condivisa di reati informatici e lo sviluppo e l'implementazione di strumenti di cooperazione internazionale.

Tale iniziativa rappresenta una risposta strutturata alla crescente diffusione del fenomeno della criminalità informatica che è divenuta componente strutturale e pervasiva della criminalità contemporanea, anche a livello transnazionale, ponendo la comunità internazionale davanti a sfide sempre più complesse.

Negli ultimi decenni, infatti, l'evoluzione della tecnologia ha radicalmente riconfigurato le modalità di commissione delle condotte illecite, alterando in modo sostanziale gli equilibri giuridici, istituzionali e geopolitici globali¹. La mutazione del crimine ha fatto però emergere un dato strutturalmente critico: la profonda asimmetria tra la rapidità del progresso tecnologico e la capacità degli ordinamenti giuridici di reagire in maniera adeguata, tempestiva ed efficace.

Le plurime lacune normative hanno infatti compromesso la possibilità

* L'articolo è stato sottoposto, in conformità al regolamento della Rivista, a referaggio "a doppio cieco".

¹ Sull'impatto della rivoluzione cibernetica sugli equilibri giuridici e sociali v. L. Picotti, *Diritto penale e tecnologie informatiche: una visione d'insieme*, in A. Cadoppi - S. Canestrari - A. Manna - M. Papa (a cura di), *Trattato di diritto penale – Cybercrime*, Milano, 2019, 33 ss.; A. Di Nicola, *Towards digital organized crime and digital sociology of organized crime*, in *Trends in Organized Crime*, 2022, 4; O. Martinu – G. McEwen, *Crime in the age of technology*, in *European law enforcement research bulletin*, 4, 2019, 23 ss.

degli Stati di esercitare un controllo effettivo sulle nuove forme di criminalità digitale, ostacolando l'attività di prevenzione, repressione e cooperazione transfrontaliera. In particolare, strumenti giuridici eterogenei e frammentati, spesso concepiti in chiave nazionale o su base regionale, si sono rivelati inadeguati a fronteggiare la dimensione globale, dinamica e tecnicamente sofisticata del *cybercrime*².

Proprio la progressiva emersione di tali lacune, nei meccanismi giuridici ed istituzionali, ha costituito una leva determinante per innescare, in sede internazionale, un processo di riflessione e dibattito orientato alla regolamentazione del cyberspazio e al contrasto coordinato della criminalità informatica.

2. Il tortuoso percorso negoziale verso un consenso globale

La rilevanza internazionale della Convenzione emerge non solo dal carattere innovativo delle disposizioni che la compongono, ma soprattutto dalla complessità del percorso negoziale che ne ha preceduto l'adozione, contraddistinto da tensioni politico-istituzionali e visioni normative contrastanti.

La risoluzione dell'Assemblea Generale delle Nazioni Unite A/RES/74/247, del 27 dicembre 2019, intitolata «*Countering the use of information and communications technologies for criminal purposes*», ha rappresentato un punto di svolta decisivo nel percorso verso la definizione della Convenzione. A tutti gli effetti, infatti, tale risoluzione può considerarsi ben oltre una semplice dichiarazione di intenti, costituendo un atto dal valore costitutivo, nel quale viene formalmente riconosciuta la necessità, condivisa dagli Stati, di definire un quadro normativo multilaterale in grado di rispondere in modo coordinato ed efficace alla crescente sofisticazione e diffusione delle minacce digitali.

Già nel preambolo della risoluzione, l'Assemblea riconosceva che le tecnologie dell'informazione e della comunicazione (ICT), pur rivestendo un ruolo chiave nello sviluppo economico ed istituzionale, vengono sempre più frequentemente strumentalizzate da soggetti criminali per condurre attività illecite di portata globale; tali condotte compromettono la stabilità delle infrastrutture e il benessere collettivo, rendendo necessario, quindi, il rafforzamento della cooperazione tra Stati, l'adozione di misure condivise e il supporto tecnico ai Paesi in via di sviluppo. Sul fronte operativo, veniva

² Per una panoramica sulle criticità legate all'uso degli strumenti informatici di ricerca della prova e sulle lacune normative, tanto a livello nazionale quanto sovranazionale, si veda M. Daniele, *La vocazione espansiva delle indagini informatiche e l'obsolescenza della legge*, in *Processo penale e giustizia*, 5, 2018; F. Spiezia, *Minaccia cibernetica e nuovi paradigmi della cooperazione giudiziaria internazionale: Il ruolo di Eurojust*, in *Sistema Penale*, 2023; più in generale, sulla c.d. "frammentazione normativa" della normativa di contrasto al *cybercrime*, cfr. anche M. Fidler, *Fragmentation of international cybercrime Law*, in *Utah law review*, 2025. Sulle nuove politiche UE in materia di *cybersecurity* e sui "rischi di incoerenza" derivanti dalle divergenze politiche, culturali, sociali, normative tra Stati membri, si veda J. Ruohonen, *The Incoherency risk in the EU's new cyber security policies*, in R. Van De Wetering, et al. (a cura di), *Disruptive innovation in a digitally connected healthy world, 23rd IFIP Conference on e-Business, e-Services, and e-Society (I3E 2024)*, 2024, 284 ss.

così istituito un comitato *ad hoc* intergovernativo incaricato di elaborare una Convenzione internazionale sul tema, sulla base degli strumenti giuridici esistenti e degli esiti delle consultazioni passate.

Come anticipato, il processo negoziale che ha condotto all'elaborazione e approvazione della Convenzione si è rivelato estremamente travagliato, tanto sotto il profilo tecnico-giuridico, quanto sotto quello politico-istituzionale³.

Fin dalla sua istituzione, il comitato *ad hoc* incaricato della redazione del testo ha dovuto operare all'interno di un quadro profondamente polarizzato, segnato da tensioni geopolitiche, divergenze normative e concettuali, oltre che da una profonda eterogeneità tra modelli giuridici di riferimento. In ordine alle scelte di criminalizzazione, ad esempio, ad alcuni Stati promotori, come la Federazione Russa e la Repubblica Popolare Cinese, favorevoli ad un impianto normativo più ampio e vincolante, si contrapponevano i Paesi occidentali, come gli Stati membri dell'Unione Europea e gli Stati Uniti, che invece manifestavano riserve sostanziali, nel timore che una definizione eccessivamente estensiva delle condotte punibili potesse tradursi in una restrizione delle libertà fondamentali.

In particolare, guardando alle relazioni presentate dai vari Stati membri durante le prime sessioni negoziali, in ordine all'estensione materiale, alla struttura e agli obiettivi della Convenzione⁴, una prima linea di frattura concerneva proprio la portata del Trattato. Da un lato, l'Unione Europea e i suoi Stati membri ne prediligevano un contenuto più restrittivo, circoscritto ai soli "*high-tech crimes*" o "*cyber-dependent crimes*", ovvero a quei crimini che non esisterebbero senza l'uso della tecnologia⁵. Su posizioni contrarie si collocava, invece, la Cina, che si faceva promotrice di un approccio più estensivo, ritenendo opportuno che nella Convenzione fossero ricompresi altresì attività criminali agevolate dalla tecnologia (*cyber-enabled crimes*), come le frodi, la pornografia, o il terrorismo⁶.

Un altro nodo centrale e critico nel confronto tra gli Stati ha riguardato la tutela dei diritti fondamentali, con particolare riguardo alla libertà di espressione e al diritto alla privacy. Sul punto, in particolare, preme se-

³ Per un'analisi del progetto di Convenzione (all'epoca in corso di negoziato), che sottolinea la crisi del rapporto tra diritto penale, sovranità e territorio dovuta alla natura transnazionale dei reati informatici, e la conseguente necessità di un trattato internazionale, si veda A. Mattarella, *La futura Convenzione Onu sul cybercrime e il contrasto alle nuove forme di criminalità informatica*, in *Sistema Penale*, 3, 2022, 41 ss.; una prospettiva simile, che affronta le obiezioni e le riserve espresse dai diversi Stati, è offerta da M. Watney, *The legal pitfalls to ratification of the United Nations Convention against cybercrime*, in *The Proceedings of the 24th European conference on cyber warfare and security*, 24(1), 2025, 713 ss.

⁴ Cfr. Un General Assembly, *Ad hoc Committee to elaborate a comprehensive international Convention on countering the use of information and communications technologies for criminal purposes*, in *unodc.org*, 2022.

⁵ Cfr. United Nations Office on drugs and crime, *Contribution from the European Union and its member states, preparation for the first session of the United Nations ad hoc Committee to elaborate a Convention on countering the use of information and communications technologies for criminal purposes taking place from 17-28 january 2022*, in *unodc.org*, 2 novembre 2021.

⁶ Cfr. United Nations Office on drugs and crime, *China's suggestions on the scope, objectives and structure (elements) of the United Nations Convention on countering the use of ICTs for criminal purposes*, in *unodc.org*, 5 novembre 2021.

gnalare la posizione europea – e sulla stessa scia anche quella del Regno Unito – favorevole ad un’impostazione garantista, secondo cui l’impiego di disposizioni vaghe e la criminalizzazione di condotte non chiaramente definite avrebbe rischiato di interferire in modo indebito e sproporzionato con la libertà di espressione⁷.

Sul fronte processuale, la questione dell’accesso e della regolamentazione delle prove digitali ha costituito ulteriore punto di attrito⁸.

Si trattava di un tema centrale, presupposto essenziale per l’efficacia dell’azione penale in materia di *cybercrime*, se si considera che, in occasione della commissione di questa tipologia di reati, le informazioni digitali vengono archiviate su server situati in giurisdizioni terze a causa dell’architettura decentralizzata dei servizi *cloud* e della globalizzazione dell’infrastruttura digitale⁹.

Ebbene, individuare un punto di equilibrio tra sovranità statale, esigenze investigative e tutela dei diritti fondamentali si è rivelato estremamente difficile. Da un lato, gli Stati Uniti sottolineavano l’importanza di meccanismi flessibili e semplificati per l’acquisizione di prove, non solo legate ai *cybercrimes* in senso stretto, ma anche ai reati commessi tramite tecnologie e, contestualmente, chiedevano che i singoli Stati si dotassero di strumenti procedurali per l’agile conservazione e condivisione di prove, sempre nel rispetto del *due process* e dei diritti fondamentali¹⁰. Dall’altro lato, Paesi come la Cina e la Russia insistevano, a tutela della sovranità statale e del principio di non interferenza, nel subordinare l’accesso ai dati a forme di cooperazione giudiziaria tradizionale¹¹.

Sul punto, la proposta europea, a tutela dei dati personali e dei diritti umani, pur condividendo la necessità di migliorare l’accesso ai dati per fini investigativi, esprimeva riserve verso qualsiasi disposizione che consentisse a

⁷ Cfr. United Nations Office on drugs and crime, *Contribution from the European Union and its member states*, cit.; in tema di diritto penale e libertà di espressione in internet, si veda il contributo di A. Costantini, *Istanze di criminalizzazione delle fake news al confine tra tutela penale della verità e repressione del dissenso*, in *Diritto penale contemporaneo*, 2, 2019, 60 ss. Sul rischio che le nuove leggi contro il *cybercrime* possano limitare la libertà di espressione per la loro formulazione eccessivamente vaga ed ampia, cfr. Council of Europe, *Freedom of expression within the context of action on cybercrime – practical considerations*, in *coe.int*, 10 dicembre 2023.

⁸ Per un’analisi delle problematiche attuali in materia di perquisizioni e sequestri informatici, conservazione di dati, intelligenza artificiale, rilevanza della prova digitale in contesti di processo penale 2.0, si veda S. De Flammineis, *Le sfide della prova digitale: sequestri, chat, processo penale telematico e intelligenza artificiale*, in *Sistema Penale*, 2024.

⁹ Sul tema del *locus commissi delicti* in materia di *cybercrimes*, cfr. R. Flor, *La legge penale nello spazio, fra evoluzione tecnologica e difficoltà applicative*, in A. Cadoppi - S. Canestrari - A. Manna - M. Papa (a cura di), *Cybercrime - Diritto e procedura penale dell’informatica*, 2023, 154 ss.; per una panoramica sull’impatto della rivoluzione cibernetica sulle categorie fondamentali del diritto penale (condotta, evento, *locus commissi delicti*) si veda anche L. Picotti, *Cybercrime e diritto penale*, in C. Parodi - V. Sellaroli (a cura di), *Diritto Penale dell’Informatica. Reati della rete e sulla rete*, 2020, 709 ss.; ancora, cfr. A. Mattarella, *Diritto penale e nuove tecnologie: dalla Convenzione ONU contro i reati informatici alle sfide dell’intelligenza artificiale*, in *Diritto penale e processo*, 2, 2025, 250 ss.

¹⁰ Cfr. United Nations Office on drugs and crime, *Submissions from member states related to the first session of the ad hoc committee*, United States, in *unodc.org*, 28 ottobre 2021.

¹¹ Cfr. United Nations Office on drugs and crime, *China’s suggestions on the scope*, cit.; United Nations Office on drugs and crime, *Submissions from member states related to the first session of the ad hoc committee*, Russian Federation, in *unodc.org*, 27 luglio 2021.

uno Stato di ottenere dati da un *provider* estero senza un adeguato controllo giurisdizionale o senza garanzie procedurali per i soggetti interessati¹². Dopo intensi dibattiti, svoltisi nell'arco di numerose sessioni tra il 2021 e i primi mesi del 2024, il testo definitivo della Convenzione è stato adottato dall'Assemblea Generale il 24 dicembre 2024 e sarà aperto alla firma dal 25 ottobre 2025 fino a dicembre 2026, per divenire vincolante solo con il raggiungimento della quarantesima ratifica.

3. Architettura e profili innovativi della Convenzione ONU

Quanto alla struttura, la Convenzione si articola in un preambolo e nove capitoli¹³.

Il primo capitolo pone le fondamenta della Convenzione, in quanto, oltre a definire i primari obiettivi e l'ambito di applicazione della stessa, evidenzia, da un lato, le regole generali applicabili all'intera Convenzione, e stabilisce, dall'altro lato, una terminologia comune, che traccia le basi concettuali e operative per l'applicazione uniforme delle norme nei diversi ordinamenti giuridici.

Dalla lettura delle preliminari disposizioni emergono immediatamente i quattro pilastri fondamentali della Convenzione: cooperazione internazionale, assistenza tecnica, armonizzazione legale e collaborazione processuale¹⁴.

Tali concetti sono infatti sin da subito messi in evidenza negli obiettivi e nell'ambito di applicazione della Convenzione, che si propone di prevenire e combattere i crimini informatici, recuperare i proventi di tali reati e rafforzare la cooperazione internazionale, anche per quanto concerne la raccolta e lo scambio di prove digitali relative a reati gravi.

Viene ribadito, inoltre, con una clausola generale conclusiva, l'obbligo per gli Stati membri di rispettare i diritti umani e le libertà fondamentali, a tutela altresì della sovranità nazionale e dell'integrità territoriale.

Su queste fondamenta si innesta il secondo capitolo, intitolato «Criminalizzazione», che pone le basi per una armonizzazione legislativa¹⁵, individuando con precisione una gamma articolata di reati che gli Stati hanno l'obbligo

¹² Cfr. United Nations Office on drugs and crime, *Written contribution from the European Union and its member states for a United Nations Convention on countering the use of information and communications technologies for criminal purposes*, in *unodc.org*, 6 aprile 2022.

¹³ Per un'analisi della struttura e dei capitoli che compongono la Convenzione cfr. anche il sito web di United Nations Office on drugs and crime in *unodc.org*.

¹⁴ Sulla cooperazione internazionale, MLA, estradizione e *challenge* procedurali nei crimini informatici, v. S. Smith - P. Torres - J. Mitchell - R. Greene, *Cybercrime and international legal cooperation*, in *Journal of international commercial law and technology*, 2, 2021, 22 ss.; anche G. Dalia, *La natura transnazionale della digital evidence tra richiesta di cooperazione e pretese di sovranità: un equilibrio necessario per il contrasto alle nuove forme di criminalità*, in *i-lex. Rivista di scienze giuridiche, scienze cognitive e intelligenza artificiale*, 16(1), 2023, 37 ss.

¹⁵ Valorizza tali aspetti anche Glen Prichard, Capo della Sezione «Cybercrime e riciclaggio di denaro» dell'Ufficio delle Nazioni Unite contro la droga e il crimine, durante l'intervento alla 13° Cybercrime Conference: cfr. *La nuova Convenzione Onu contro il cybercrime*, in *ictsecuritymagazine.com*, 2025.

di introdurre nei propri ordinamenti penali. Si tratta, in particolare, di reati informatici, sia “*cyber-dependent*”, ovvero dipendenti dalle tecnologie, sia, seppur in numero ridotto, “*cyber-enabled*”, ovvero agevolati dalle stesse¹⁶. Rispetto a quest’ultima categoria, degna di nota è l’introduzione di nuove condotte, mai incluse prima in trattati internazionali, come la diffusione non consensuale di immagini intime o l’adescamento online di minori.

Con il capitolo seguente, il terzo, si tenta di risolvere la questione, particolarmente spinosa, delle competenze territoriali e dei conflitti di giurisdizione in caso di reati digitali, che non di rado trascendono i confini nazionali. In merito, si richiede agli Stati di stabilire la propria giurisdizione almeno nei casi in cui il reato sia commesso sul proprio territorio, da o a danno di un cittadino nazionale, quando lo Stato ne sia direttamente colpito, nonché laddove l’autore sia ancora presente sul territorio dello Stato. In caso di giurisdizioni concorrenti, però, si introducono dei meccanismi obbligatori di consultazione finalizzati a prevenire o risolvere i conflitti.

A questo punto, la Convenzione si dedica all’aspetto operativo-investigativo con il capitolo quarto «Misure procedurali e applicazione della legge», che introduce strumenti specifici mirati all’identificazione, alla raccolta e alla conservazione delle prove digitali.

Nel bilanciamento tra efficienza investigativa e protezione dei diritti fondamentali, viene imposta l’adozione di meccanismi di preservazione dei dati, di perquisizione e sequestro digitale, di emissione di ordini di produzione per i *provider*, di intercettazione di traffico in tempo reale¹⁷. Tali meccanismi vengono ad ogni modo accompagnati da rigorose garanzie procedurali, quali la necessità del controllo giudiziario, il rispetto del principio di proporzionalità, l’obbligo di motivazione e il diritto al ricorso.

L’attuazione delle disposizioni procedurali appena viste avrebbe rischiato di rimanere solo sulla carta se non fosse stato istituito anche un articolato sistema di cooperazione internazionale, che viene per l’appunto dettagliato nel capitolo quinto, dove si delinea un *framework* multilivello di assistenza reciproca tra Stati in materia di estradizione, trasferimento dei procedimenti, indagini congiunte e scambio di prove digitali, non solo nel caso in cui i reati siano strettamente informatici, ma anche semplicemente nelle ipotesi in cui rientrino nella categoria dei reati gravi. Occorre segnalare, all’interno di questo capitolo, l’istituzione di una rete di contatto rapido

¹⁶ La distinzione tra *cyber-dependent* e *cyber-enabled crimes* risiede nel diverso grado di dipendenza dal mezzo informatico: i primi sono reati che esistono unicamente grazie alle tecnologie digitali (ad es. l’accesso abusivo a un sistema informatico o l’interferenza con dati e sistemi), mentre i secondi sono reati “tradizionali” che trovano nel cyberspazio un canale di amplificazione o facilitazione (ad es. il traffico di droga, il riciclaggio di denaro attraverso criptovalute o la diffusione di materiale di abuso sessuale su minori); in merito a tale distinzione cfr. United Nations Office on drugs and crime, *Digest of cyber organized crime*, Vienna, 2, 2022, 36 ss.; European Parliament, *Understanding cybercrime*, in *European Parliament research service*, 2024; S. Cross - S. Hirle - M.A. Lim, *Cybercrime strategy guidebook*, Interpol, 2021; cfr. O. Martinu - G. McEwen, *Crime in the age of technology*, cit., 23 ss.

¹⁷ In tema di bilanciamento tra strumenti di investigazione e principi fondamentali, cfr. S. Signorato, *Il sequestro di dispositivi e informazioni digitali*, in *Sistema Penale*, 2025, cfr. anche J. Della Torre, *Spunti di riflessione sulla proposta di legge in materia di sequestro di dispositivi e di dati*, in *Sistema Penale*, 2025, sulla proposta di legge C. 1822, approvata dal Senato, recante «Modifiche al Codice di procedura penale in materia di sequestro di dispositivi, sistemi informativi o telematici o memorie digitali».

24h, pensata proprio per facilitare la comunicazione in tempo quasi reale tra autorità competenti.

Una volta definito il quadro repressivo, con il capitolo sesto si introduce una prospettiva preventiva, mirata proprio alla riduzione della diffusione del *cyber crime*, tramite la promozione della consapevolezza pubblica, il rafforzamento delle competenze istituzionali e il coinvolgimento di tutti gli attori sociali.

A completamento di queste misure, il settimo capitolo evidenzia l'importanza di colmare le disparità tecnologiche tra gli Stati attraverso il supporto nella formazione degli operatori, nell'analisi forense e nell'uso delle tecnologie investigative, e il continuo monitoraggio delle tendenze criminali, nonché il reciproco scambio di buone pratiche.

Infine, nel capitolo sulle procedure di attuazione della Convenzione, si prevede l'istituzione della Conferenza delle parti, quale organismo permanente, seppur flessibile, incaricato di valutare l'efficacia delle misure adottate dagli Stati, raccogliere dati, avanzare proposte di modifiche o integrazioni.

In chiusura, le disposizioni finali mirano a disciplinare gli aspetti tecnici legati all'entrata in vigore del Trattato, all'adesione e alla possibilità di recesso.

4. La Convenzione ONU e la Convenzione di Budapest: due modelli a confronto

Esaurita l'analisi della struttura, dei contenuti e degli obiettivi perseguiti dalla nuova Convenzione delle Nazioni Unite sulla criminalità informatica, si ritiene utile, una riflessione comparativa con la Convenzione di Budapest, un altro fondamentale trattato internazionale sulla lotta contro la criminalità informatica, adottato nel 2001 in seno al Consiglio d'Europa¹⁸. Tale raffronto appare opportuno non solo per la funzione pionieristica che quest'ultima Convenzione ha svolto nella costruzione di un *corpus* normativo sul *cybercrime*, ma anche per il fatto che la recente Convenzione ONU è stata espressamente concepita, dai suoi promotori, come un'integrazione e una possibile estensione, in chiave globale, della Convenzione di Budapest. La comparazione tra i due strumenti consente dunque di mettere in luce gli elementi di continuità normativa e le differenze sostanziali sotto il profilo della legittimazione negoziale, della portata giuridica, delle tecniche di criminalizzazione e dei modelli di cooperazione internazionale adottati.

La Convenzione di Budapest rappresenta il primo e più longevo trattato internazionale vincolante dedicato alla lotta contro la criminalità informatica. È ratificata da 76 Stati, membri e non del Consiglio d'Europa, e si fonda su un impianto normativo mirato ad armonizzare le normative penali nazionali in materia di reati informatici, affiancato da un apparato

¹⁸ Propongono un'analisi comparata delle due Convenzioni anche A. Kazakova - S. Teleanu - B. Kovac - F. Hamza, *Comparative analysis: the Budapest Convention vs the UN Convention against cybercrime*, in *Digital watch observatory*, 22 ottobre 2024; T. Tropina, "This is not a human rights convention!": The perils of overlooking human rights in the UN cybercrime treaty, in *Journal of cyber policy*, 9(2), 2024.

di disposizioni procedurali sulla cooperazione giudiziaria e sullo scambio transfrontaliero di prove digitali¹⁹.

Il corpus convenzionale si è arricchito nel tempo di due protocolli addizionali: il primo, del 2003, per il contrasto alla xenofobia e al razzismo online; il secondo, del 2022 – ancora in fase di ratifica – volto al potenziamento della cooperazione rafforzata e delle procedure per la circolazione delle prove digitali²⁰.

Benché entrambe le Convenzioni condividano l'obiettivo generale di contrasto alla criminalità informatica, esse si differenziano tanto per l'ambito di applicazione, quanto per le finalità operative.

La Convenzione di Budapest ha un'impostazione prevalentemente repressiva e settoriale, incentrata sulla criminalizzazione di specifiche condotte (ad esempio l'accesso illecito, le interferenze su dati e sistemi, le frodi informatiche, la pornografia minorile) e sull'introduzione di strumenti investigativi e cooperativi volti a garantire l'efficacia transnazionale dell'azione penale. Il suo impianto normativo è tecnicamente rigoroso, incentrato sui reati cibernetici in senso stretto, e mira a stabilire un quadro giuridico comune tra gli Stati, pur riconoscendo a questi ultimi margini di adattamento alle proprie peculiarità giuridiche.

Di contro, la Convenzione ONU adotta un approccio integrato e multidimensionale, in cui la repressione penale rappresenta solo una delle componenti di un articolato progetto di *governance* della sicurezza digitale. Essa include misure di prevenzione, programmi di rafforzamento di *capacity building*, assistenza tecnica e tutela delle infrastrutture critiche, con particolare attenzione alle esigenze dei Paesi in via di sviluppo.

Tale impostazione riflette una concezione più ampia della *governance* del cyberspazio, non circoscritta alla sola repressione dei reati, ma volta alla creazione di condizioni strutturali per una sicurezza informatica globale e condivisa. Inoltre, la Convenzione ONU si distingue per l'esplicita valorizzazione della sovranità statale, quale parametro fondamentale nell'attuazione delle disposizioni convenzionali, assente nella Convenzione di Budapest.

Sotto il profilo della criminalizzazione, la Convenzione ONU prevede poi un significativo ampliamento del novero delle condotte penalmente sanzionabili, includendo nel proprio catalogo incriminatorio una serie di fatti-

¹⁹ Per una panoramica sulla normativa di contrasto al *cybercrime*, dalla Convenzione di Budapest alla Convenzione ONU cfr. anche M. Fidler, *Fragmentation of international cybercrime law*, cit.; sulla disciplina della cooperazione internazionale introdotta dalla Convenzione di Budapest si veda A. Balsamo, *Spazio virtuale e processo penale: la nuova Convenzione ONU sul cybercrime*, in *Diritto penale e processo*, 2, 2025, 240 ss.; E. Colombo, *La cooperazione internazionale nella prevenzione e la lotta alla criminalità informatica: dalla Convenzione di Budapest alle disposizioni nazionali*, in *Cyberspazio e diritto*, 3-4, 2009, 285 ss.; anche F. Cajani, *La Convenzione di Budapest nell'insostenibile salto all'indietro del Legislatore italiano, ovvero: quello che le norme non dicono*, in *Cyberspazio e diritto*, 1, 2010, 185 ss.

²⁰ Sulle novità introdotte dal Secondo Protocollo aggiuntivo alla Convenzione di Budapest e l'impatto con il diritto alla riservatezza, v. L. Di Crescenzo, *Conservazione e circolazione delle informazioni per fini investigativi condizioni e garanzie in materia di protezione dei dati personali*, in *Rivista di scienze giuridiche, scienze cognitive ed intelligenza artificiale*, 15(2), 2022, 27 ss.; anche M. Buccarella, *Il secondo protocollo addizionale alla Convenzione di Budapest e le nuove frontiere della cooperazione internazionale in ambito digitale. Quali rischi per la protezione dei dati personali nell'Unione europea?*, in *Associazione italiana studiosi di diritto dell'Unione europea*, 8, 2023, 184 ss.

specie che riflettono un orientamento normativo più attento ad intercettare le minacce digitali emergenti. Tra le nuove incriminazioni si annoverano la già menzionata criminalizzazione dell'adescamento online di minori, la diffusione non consensuale di immagini intime, nonché il riciclaggio di proventi derivanti da reati informatici.

Particolarmente indicativa è la disposizione di cui all'art. 15 della Convenzione ONU che, a differenza dell'art. 9 della Convenzione di Budapest - riguardante il possesso e la diffusione di materiale pedopornografico - estende la sfera della rilevanza penale anche a condotte prodromiche a tale reato, come la sollecitazione, il grooming e la predisposizione di incontri a fini illeciti. Si tratta di un'evoluzione significativa, che evidenzia la volontà di intervenire, in via preventiva, già nella fase di strutturazione del contesto criminogeno, sanzionando condotte che, pur non producendo un danno immediato, ne costituiscono l'antecedente logico e funzionale.

Parallelamente, la scelta di escludere dal novero dei reati le violazioni in materia di proprietà intellettuale e diritto d'autore, specificamente trattati invece nel corpus normativo della Convenzione di Budapest, riflette una diversa gerarchizzazione dei beni giuridici meritevoli di protezione. Pare che la Convenzione ONU tenda a privilegiare, almeno per il momento, una selezione più essenziale e universalmente condivisibile delle condotte offensive, rinunciando ad alcune fattispecie che avrebbero potuto ostacolare il consenso multilaterale.

Proprio alla luce di questo intento di intercettare con maggiore efficacia le minacce emergenti dall'uso delle tecnologie digitali, anche sul piano concettuale, la Convenzione ONU sceglie di adottare un lessico giuridico più ampio e tecnologicamente aggiornato rispetto al modello europeo: ad esempio, contrariamente alla Convenzione di Budapest che parla di "dati e sistemi informatici", il testo ONU predilige una terminologia più ampia, utilizzando termini come "sistemi ICT" e "dati elettronici", in modo da estendere l'ambito applicativo delle norme ad un ventaglio più vasto di dispositivi ed infrastrutture.

Infine, la recente Convenzione circoscrive l'attivazione degli strumenti di cooperazione internazionale ai soli reati gravi, intesi come quelli sanzionati con una pena detentiva non inferiore a quattro anni, restringendo l'ambito operativo della cooperazione rispetto alla Convenzione di Budapest, che invece si applica anche a fattispecie meno gravi, purché restino coinvolte le prove digitali.

5. Le perplessità delle organizzazioni internazionali alla ricerca di un equilibrio tra cooperazione internazionale e tutela dei diritti

L'adozione della Convenzione delle Nazioni Unite contro la criminalità informatica segna una svolta epocale nel quadro della cooperazione internazionale per la regolamentazione giuridica del cyberspazio e della governance digitale globale.

Per la prima volta, un organismo multilaterale di portata universale si è dotato di uno strumento giuridico vincolante interamente concepito per il

contrasto alla criminalità informatica.

In tale contesto, la Convenzione non si limita ad assumere i tratti di un testo normativo innovativo, ma si configura come una manifestazione concreta della volontà degli Stati membri di affrontare, con strumenti armonizzati, l'evoluzione pervasiva delle minacce digitali²¹.

Uno dei meriti più rilevanti della Convenzione consiste nell'aver istituzionalizzato un'architettura multilivello di cooperazione internazionale, che prevede meccanismi avanzati di assistenza giudiziaria, di scambio di informazioni elettroniche e di conservazione delle prove digitali.

Inoltre, meritevole di rilievo è l'approccio estensivo dell'ambito oggettivo di applicazione della Convenzione che ha il pregio di non limitarsi ai reati strettamente "*cyber-dependent*", ma di ricomprendere nel proprio ambito applicativo anche alcuni crimini "*cyber-enabled*", offrendo una copertura normativa flessibile, che rispecchia la natura ibrida delle condotte criminali contemporanee.

Altrettanto significativa è l'impostazione inclusiva e partecipativa che ha contraddistinto l'intero processo negoziale, durante il quale è stato garantito uno spazio di interlocuzione strutturata alla società civile, agli esperti accademici e agli attori non governativi²².

Non va altresì sottovalutato il valore politico della Convenzione, che, malgrado l'evidente eterogeneità degli interessi geopolitici in gioco, ha raccolto un ampio consenso trasversale. Tra le novità degne di nota si annovera, ad esempio, l'introduzione, sia pure parziale, di una prospettiva di genere tramite l'esplicito riconoscimento della dimensione sessuata del *cybercrime*, la menzione della violenza di genere in ambienti digitali e l'invito al *mainstreaming* di genere in sede di attuazione²³.

Parallelamente, la previsione di strumenti per il *capacity building* e il rafforzamento istituzionale nei Paesi meno attrezzati, mediante programmi di assistenza tecnica, formazione e scambio di buone pratiche, evidenzia l'obiettivo dichiarato di ridurre il divario globale in materia di sicurezza cibernetica.

Ciononostante, accanto ai profili positivi sin qui delineati, una riflessione critica e rigorosa impone di considerare le numerose riserve espresse da una pluralità di attori della comunità internazionale e, in particolare, da

²¹ Cfr. J. Hakmeh, *The UN convention on cybercrime: a milestone in cybercrime cooperation?*, in *Journal of cyber policy*, 9(2), 2024, 125 ss.

²² Sul punto, v. I.A. Barber – S. Kumar, *Learning from the ground up: lessons from civil society engagement in addressing the human rights implications of cybercrime legislation*, in *Journal of cyber policy*, 9(2), 2024, 131 ss.

²³ Sulle questioni di genere, cfr. J. Hakmeh, *The UN convention on cybercrime: a milestone in cybercrime cooperation?*, cit., che, trattando la "*gender dimension*", definisce il "*gender mainstreaming*" come il processo con cui si assicura che la prospettiva di genere e l'uguaglianza tra uomini e donne siano integrate nelle politiche, nei programmi e nelle pratiche istituzionali; sul *gender mainstreaming* si veda anche Chatham House Cyber Policy team, *Gender mainstreaming and the proposed cybercrime convention: Commentary on the consolidated draft*, in *chathamhouse.org*, 2022; infine, un'analisi dell'approccio femminista nelle leggi sul *cybercrime*, incluso l'esame sul punto dell'*UN Cybercrime Convention* già nella fase di negoziato è offerta da Centre for feminist foreign policy – P. Pavlova, *Intersectional feminist perspectives on cybercrime law toward an effective and human rights-respecting implementation of the United Nations Cybercrime Convention*, in *centreforfeministforeignpolicy.org*, 2025.

parte delle organizzazioni impegnate nella tutela dei diritti umani. Queste ultime hanno sollevato preoccupazioni in merito a talune disposizioni del Convenzione, ritenute ambigue o eccessivamente permissive, evidenziando il rischio che, in assenza di garanzie sostanziali e vincolanti, la Convenzione potesse essere strumentalizzata per giustificare pratiche di sorveglianza abusiva, violazioni alla privacy ed arbitrarie limitazioni alla libertà di espressione.

Tali preoccupazioni sono state formalizzate in un documento congiunto, una lettera pubblica sottoscritta da una serie di organizzazioni internazionali per i diritti umani, che ha assunto il valore di un appello politico e giuridico inequivocabile²⁴. Si chiede, infatti, agli Stati membri di non procedere all'approvazione della Convenzione nella sua attuale formulazione, evidenziandone i rischi per l'equilibrio tra sicurezza e libertà fondamentali. In generale, le preoccupazioni espresse attengono principalmente alla struttura stessa del testo normativo, alla vaghezza di talune disposizioni e all'assenza di garanzie vincolanti; si teme che la Convenzione possa divenire un veicolo di legalizzazione internazionale per prassi autoritarie, legittimanti abusi e compressioni delle libertà civili.

Una delle principali criticità individuate risiede nell'ambito di applicazione della Convenzione, che appare fin dall'inizio eccessivamente esteso e privo di delimitazioni chiare.

In particolare, l'art. 23 impone agli Stati l'obbligo di raccogliere prove digitali non solo per reati informatici, ma per una vasta gamma di crimini comuni, indipendentemente dal fatto che essi coinvolgano o meno tecnologie dell'informazione e della comunicazione. Similmente, l'art. 35, senza prevedere salvaguardie procedurali né limiti contenutistici, obbliga alla cooperazione transfrontaliera per "qualsiasi crimine grave", definito genericamente come reato punibile con una pena minima di quattro anni. A tale impostazione estensiva contribuisce altresì l'art. 4, che criminalizza una serie di condotte in termini vaghi e indeterminati, contravvenendo al principio di legalità penale, che impone una descrizione chiara, precisa e tassativa delle fattispecie incriminatrici²⁵.

In tal modo, la Convenzione diviene – secondo le organizzazioni appellanti – un potenziale strumento di sorveglianza globale, che potrebbe essere utilizzato da regimi autoritari per monitorare oppositori politici, giornalisti, attivisti e minoranze vulnerabili, giustificando tali pratiche sotto la veste della cooperazione internazionale in materia penale²⁶.

In secondo luogo, sempre sul piano dei diritti fondamentali, la Convenzione sarebbe altresì caratterizzata da una debolezza strutturale delle tutele procedurali.

²⁴ *Joint letter to the European Union and its member States concerning the United Nations Cybercrime Convention*, in *brn.org*, 21 ottobre 2024, firmata da Access Now, Amnesty International, ApTI, Romania Asli Telli, WISER at Wits University, Centre for Feminist Foreign Policy (CFFP), Chaos Computer Club, CyberPeace Institute, Digitalcourage, Douwe Korff, Emeritus Professor of international law, London Metropolitan University Electronic Frontier Foundation (EFF), Electronic Frontier Norway, Electronic Privacy Information Center (EPIC), European Digital Rights (EDRi).

²⁵ *Ibid.*

²⁶ Ivi, 1.

In proposito, l'art. 24, dedicato alle salvaguardie, si limita a evocare genericamente il principio di proporzionalità, omettendo riferimenti vincolanti ai principi di legalità, necessità, e non discriminazione, pilastri fondamentali del diritto internazionale dei diritti umani. Peraltro, si attribuisce direttamente agli Stati membri la responsabilità di implementare tali salvaguardie, senza una cornice comune o standard minimi obbligatori, il che produce un effetto di frammentazione normativa e indebolisce la protezione effettiva dei diritti nei contesti più vulnerabili²⁷.

Ciò comporta, anche secondo parte della dottrina, il rischio di una tutela “a geometria variabile”, dove le garanzie dipendono dalla sensibilità giuridica del singolo Stato, favorendo la disparità di trattamento e la possibilità di abusi²⁸. Tale debolezza è stata esplicitamente sottolineata anche durante i negoziati, nei quali alcuni delegati, come il rappresentante iraniano, hanno criticamente dichiarato che «*This is not a human rights convention!*», suggerendo che l'obiettivo primario della Convenzione non fosse la tutela delle libertà, bensì l'efficienza repressiva²⁹.

Analoga problematicità si riscontra nella struttura delle misure investigative previste – come l'intercettazione di dati in tempo reale, il sequestro di dispositivi o la raccolta di metadati – che non prevedono limiti specifici né condizioni stringenti, divenendo dunque applicabili a un'ampia gamma di reati.

In questo quadro, l'assenza di esplicite clausole esimenti a tutela della attività giornalistica, del *whistleblowing*, della ricerca sulla sicurezza informatica e della denuncia civica di vulnerabilità, espone soggetti come giornalisti, attivisti, ricercatori e segnalatori civici a rischio di incriminazioni arbitrarie; inoltre, la mancata definizione del concetto di “intento doloso” contribuisce a questa incertezza, lasciando margine alla criminalizzazione di condotte esercitate nell'interesse pubblico in assenza di criteri normativi chiari e proporzionati³⁰.

Per quanto concerne la disciplina sulla responsabilità degli intermediari digitali, la Convenzione introduce elementi di significativa ambiguità interpretativa. Come infatti specificato all'interno della lettera menzionata, da un lato, l'art. 18 configura una forma di responsabilità oggettiva per la

²⁷ Ivi, 2.

²⁸ Per un'esauritiva analisi critica dell'impatto e dei rischi dell'obbligo di incriminazione, della delimitazione della competenza giurisdizionale tra gli Stati e dei pervasivi meccanismi di cooperazione giudiziaria internazionale sulla protezione dei diritti umani e delle libertà fondamentali cfr. M. Dimetto, *Convenzione delle Nazioni Unite contro il cybercrime e tutela dei diritti umani: influenze europee sullo scenario internazionale*, in *Freedom, security & justice*, 2, 2025, 108 ss.; ancora, sul punto, si veda K. Bannelier, *Risks and opportunities of the UN Cybercrime Convention for the UNDOC & the fight against transnational organised crime: a first assessment*, in *transnational criminal law review*, 2025, che manifesta la preoccupazione che la Convenzione possa essere sfruttata, in sede di implementazione, dai regimi autoritari; tali preoccupazioni vengono sollevate anche da alcune organizzazioni per i diritti umani, come Article 19, *Article 19's analysis of the UN Convention against cybercrime*, in *article19.org*, 2024, o come nell'analisi critica delle ONG più autorevoli sulla difesa dei diritti digitali, Electronic Frontier Foundation, *The UN Cybercrime draft convention remains too flawed to adopt*, in *eff.org*, 2024.

²⁹ Sul punto, si vedano altresì le perplessità sollevate da T. Tropina, “*This is not a human rights convention!*”: *The perils of overlooking human rights in the UN cybercrime treaty*, cit.

³⁰ Cfr. *Joint letter to the European Union and its member states*, cit., 3.

partecipazione a reati informatici, a prescindere dall'esistenza di un intento doloso, e, dall'altro lato, l'art. 19, in evidente contraddizione, richiede invece la consapevolezza dell'infrazione. Tale disallineamento può produrre conseguenze rilevanti in termini applicativi, incentivando le piattaforme digitali ad adottare pratiche di moderazione preventiva eccessivamente restrittive al fine di evitare implicazioni giuridiche, ma che nel lungo termine si tradurrebbero in restrizioni della libertà di espressione online. Si evidenzia, peraltro, una chiara dissonanza con il quadro normativo europeo e, in particolare, con il *Digital Services Act*, che riconosce una forma di immunità per gli intermediari rispetto ai contenuti generati dagli utenti, subordinata proprio alla conoscenza effettiva dell'illecito³¹.

Parallelamente, la prospettiva di genere risulta ampiamente sottorappresentata nel corpo della Convenzione, nonostante il preambolo richiami formalmente l'importanza dell'integrazione di una prospettiva di genere (*"gender mainstreaming"*). Le uniche disposizioni (artt. 34 e 53, lett. h) esplicitamente riferite alla questione, infatti, appaiono marginali e prive di effettività giuridica. Infine, nella lettera si precisa come, a testimonianza di una resistenza politica alla piena inclusione dei principi di equità, parità e inclusività, le proposte avanzate dalla società civile per rafforzare la dimensione intersezionale della sicurezza informatica non siano state recepite nel testo finale³².

Particolari criticità vengono poi segnalate in relazione alle disposizioni relative alla cooperazione giudiziaria internazionale e al trasferimento transfrontaliero di dati.

In particolare, l'art. 40 della Convenzione introduce un obbligo generale di assistenza giudiziaria «nel modo più ampio possibile», formula vaga, volutamente indeterminata, che solleva interrogativi in merito alla sua compatibilità con normative nazionali ed europee in materia di protezione dei dati, come il regolamento generale sulla protezione dei dati (GDPR) e la direttiva UE sull'applicazione della legge (c.d. Direttiva Enforcement).

Ulteriori perplessità si riscontrano negli artt. 36 e 47, che autorizzano la conclusione di accordi bilaterali e multilaterali finalizzati a facilitare il trasferimento di dati, anche verso Paesi che non rispettano standard minimi di tutela dei diritti fondamentali, aprendo la possibilità a flussi informativi verso ordinamenti privi di un quadro di garanzie comparabili a quelle europee.

Inoltre, la Convenzione incoraggia forme di cooperazione diretta tra autorità di polizia, che escludono l'intervento delle autorità giudiziarie o di controllo indipendenti: si pensi, ad esempio, all'istituzione della rete 24/7 che, sebbene concepita per agevolare lo scambio rapido di informazioni tra Stati membri, si configura come un canale parallelo a quello giudiziario, privo di garanzie procedurali sufficienti con conseguente indebolimento della trasparenza, della tracciabilità e del controllo sull'uso dei dati³³.

Sempre sul fronte della cooperazione giudiziaria, si evidenzia, infine, la mancata previsione del principio di doppia incriminazione nelle disposi-

³¹ Ivi, 2.

³² *Ibid.*

³³ Ivi, 3.

zioni della Convenzione: l'assenza di tale garanzia consentirebbe agli Stati di richiedere la cooperazione anche per reati che non sono riconosciuti come tali nell'ordinamento dello Stato richiesto, sollevando rilevanti preoccupazioni sotto il profilo della sovranità giuridica e della tutela dei principi costituzionali dei paesi democratici³⁴.

6. Conclusioni

Giunti all'epilogo di questa analisi, si impone una considerazione conclusiva sul significato e sull'impatto della Convenzione ONU sulla criminalità informatica.

È indubbio che l'adozione di tale strumento normativo rappresenti un passaggio epocale nell'ambito della regolazione multilaterale del cyberspazio. Era infatti divenuta improcrastinabile la definizione di un quadro giuridico condiviso che consentisse agli Stati di superare la frammentazione normativa in materia di *cybercrimes*, e cooperare nella prevenzione, repressione e gestione di tali crimini.

Tuttavia, accanto a tale valore strategico, emergono con chiarezza significative criticità. Le ambiguità lessicali, l'eccessiva ampiezza dell'ambito applicativo, la carenza di salvaguardie procedurali vincolanti, la debolezza delle garanzie in materia di diritti fondamentali sollevano legittime riserve circa i rischi potenziali di un uso strumentale o distorsivo delle disposizioni convenzionali.

Nel timore che uno strumento nato per contrastare il crimine digitale venga piegato a logiche di sorveglianza e repressione, alimentando pratiche autoritarie sotto la copertura della cooperazione internazionale, appare dunque necessaria una lettura prudente e critica del testo. A questo fine, sarà poi decisivo il ruolo degli attori democratici - istituzioni, società civile, mondo accademico - nel vigilare su un'attuazione della Convenzione che sia compatibile con i diritti fondamentali e con le garanzie sostanziali e procedurali previste a livello nazionale.

Il vero banco di prova sarà pertanto la capacità degli Stati di interpretare e applicare la Convenzione nel rispetto dei principi dello Stato di diritto, dei diritti umani e delle garanzie democratiche, nella costante ricerca di un equilibrio tra sicurezza e libertà.

³⁴ Sul punto, sempre T. Tropina, *"This is not a human rights convention!": The perils of overlooking human rights in the UN cybercrime treaty*, cit.

Abstract

L'articolo analizza il processo di elaborazione e i contenuti della Convenzione delle Nazioni Unite sul crimine informatico, esaminando il contesto delle sfide globali poste dalla criminalità digitale e le tensioni geopolitiche che ne hanno accompagnato i negoziati. Dopo una premessa introduttiva, il contributo ricostruisce il complesso percorso negoziale che ha condotto all'adozione dello strumento convenzionale, evidenziando le divergenze tra gli Stati in merito a sovranità, sicurezza e diritti fondamentali. Vengono quindi esaminati l'architettura della Convenzione e i suoi profili innovativi, con particolare attenzione ai meccanismi di cooperazione internazionale. Non manca, poi, un confronto con la Convenzione di Budapest al fine di mettere in luce le differenze tra i due modelli normativi. Infine, l'articolo rappresenta le perplessità espresse dalle organizzazioni internazionali e dalla società civile, interrogandosi sulla capacità della Convenzione ONU di bilanciare efficacemente l'esigenza di cooperazione repressiva con la tutela dei diritti umani fondamentali.

The article analyses the drafting process and the content of the United Nations Convention on Cybercrime, examining the context of the global challenges posed by digital crime and the geopolitical tensions that accompanied its negotiation. After an introductory section, the contribution reconstructs the complex negotiation process that led to the adoption of the conventional instrument, highlighting the divergences among States with regard to sovereignty, security, and fundamental rights. It then examines the architecture of the Convention and its innovative features, with particular attention to the mechanisms of international cooperation. The article also includes a comparison with the Budapest Convention in order to highlight the differences between the two normative models. Finally, it presents the concerns expressed by international organizations and civil society, questioning the ability of the UN Convention to effectively balance the need for repressive cooperation with the protection of fundamental human rights.

Keywords

Convenzione *cybercrime* - cooperazione internazionale - frammentazione normativa - diritti fondamentali - crimine informatico

Note

Transmission of Intercepted Data Obtained in a Criminal Investigation to the Competition Authority: The Grand Chamber Judgment in the Ships Waste Oil Case*

Claudia Calabrese

European Court of Human Rights, 1 April 2025, *Ships Waste Oil Collector B.V. and Others v. the Netherlands*, app. nos. 2799/16, 3124/16, 3205/16 and 2800/16

The transmission of data collected for criminal investigations to an administrative authority for other purposes constitutes an interference with the right to respect for correspondence protected by art. 8 of the European Convention on Human Rights. In the Ships Waste Oil Collector B.V. case, however, the transfer of data was found compatible with art. 8, as it was based on data collected in a conventionally compliant manner, supported by a sufficiently accessible and predictable legal basis, and accompanied by appropriate safeguards to prevent arbitrariness, including ex post judicial review. In consideration of the predominantly economic and commercial character of the information, and the objective of safeguarding the country's economic well-being, the Grand Chamber determined that the interference was proportionate to the requirements of a democratic society founded on the rule of law.

Table of contents

1. Introduction: Fundamental Rights and Legal Persons under the ECHR. – 2. The National Proceedings. – 3. The Status of Companies as “Victims” under the European Convention on Human Rights. – 4. The Margin of Appreciation and the Protection of Strong Public Interests. – 5. When Interference Becomes a Violation: a Comparison of Two Approaches. – 6. The Importance of Art. 39 of the Dutch Judicial and Criminal Records Data Act (WJSG). – 7. Concluding Remarks.

* L'articolo è stato sottoposto, in conformità al regolamento della Rivista, a referaggio anonimo.

This Paper has received funding from the European Union's Horizon Europe Research and Innovation programme under grant agreement N°101168309.

1. Introduction: Fundamental Rights and Legal Persons under the ECHR

The protection of fundamental rights has historically centred on natural persons. Nevertheless, legal persons are increasingly recognized as holders of constitutional rights and, in some cases, of international human rights¹. The European Convention on Human Rights (ECHR) is a prime example of this approach, as its application is not limited solely to individuals². Art. 1 ECHR provides: «The High Contracting Parties shall secure to everyone within their jurisdiction the rights and freedoms defined in Section I of this Convention», thereby extending the protection of these rights and freedoms to all legal subjects within domestic legal systems. It is evident that several provisions within the Convention and its Protocols serve to reinforce this interpretation. Art. 10, para. 1, ECHR enables broadcasting and media enterprises to be subject to licensing regimes³. Art. 1, para. 1, Protocol No. 1, ECHR explicitly protects the property rights of «every natural or legal person»⁴. Art. 34 ECHR regulates the individual application procedure, extending it to «non-governmental organizations or groups of individuals»⁵.

¹ For an overview of the range of rights open to companies under the ECHR, see M. Emberland, *The Human Rights of Companies*, Oxford, 2006, 156 ff.; S.K. Ripken, *Corporations are people too: A multidimensional approach to the corporate personhood puzzle*, in *Fordham Journal of Corporate & Financial Law*, 15(1), 2009, 97 ff.; D. Rubin, *Corporate personhood: How the courts have employed bogus jurisprudence to grant corporations constitutional rights intended for individuals*, in *Quinnipiac Law Review*, 28(2), 2010, 4 ff.; M. Julicher – M. Henriques – A.A. Blai – P. Policastro, *Protection of the EU Charter for Private Legal Entities and Public Authorities? The Personal Scope of Fundamental Rights within Europe Compared*, in *Utrecht Law Review*, 15(1), 2019, 1 ff.; S. Watson, *The Corporate Legal Person*, in *Journal of Corporate Law Studies*, 19(1), 2018, 137; E. Gill-Pedro, *Proportionality and the Human Rights of Companies under the ECHR - Whose Interests Are at Stake?*, in *Nordic Journal of International Law*, 89(3), 2020, 327 ff.; L. Dopplinger, *Legal Persons as Bearers of Rights Under the ECHR*, in *University of Vienna Law Review*, 5(1), 2021, 1 ff.

Moreover, for further information on the invocation of art. 8 ECHR by legal persons, see B.A. Juliussen, *Monitoring, Governmental Data Access and the invocation of Article 8 ECHR by Legal Persons*, in *International Journal of Human Rights*, 29(7), 2025, 1251 ff.

² The reference is to art. 2, para. 1, International Covenant on Civil and Political Rights, which specifies: «Each State party undertakes to respect and to ensure to all individuals within its territory and subject to its jurisdiction the rights recognized in the present Covenant, without distinction of any kind, such as race, color, sex, language, religion, political or other opinion, national or social origin, property, birth or other status». See also the American Convention on Human Rights (ACHR), for the purposes of which, in accordance with art. 1, para. 2, “person” means “every human being”.

³ Art. 10, para. 1, ECHR provides: «Everyone has the right to freedom of expression. This right shall include freedom to hold opinions and to receive and impart information and ideas without interference by public authority and regardless of frontiers. This Article shall not prevent States from requiring the licensing of broadcasting, television or cinema enterprises».

⁴ Art. 1, para. 1, Protocol No. 1, ECHR states: «Every natural or legal person is entitled to the peaceful enjoyment of his possessions. No one shall be deprived of his possessions except in the public interest and subject to the conditions provided for by law and by the general principles of international law».

⁵ Art. 34 ECHR reads: «The Court may receive applications from any person, non-

Nevertheless, the European Court of Human Rights (ECtHR) has played a decisive role beyond these explicit provisions by progressively broadening the scope of the Convention. The Court has recognized that legal entities may also be holders of rights not expressly attributed to them in the text, provided they meet the criteria to qualify as “victims” or can credibly claim such status.

It is important to note that not all Convention rights are automatically enforceable by legal persons; only those rights that can be meaningfully exercised, such as property rights, the right to a fair trial, freedom of expression, and the protection of correspondence and domicile, are typically relevant.

The Grand Chamber judgment of 1 April 2025 in the *Ships Waste Oil Collector B.V.*⁶ addresses the applicability of art. 8 ECHR to legal persons, particularly regarding the protection of commercial data transmitted between public authorities.

The applicants are six limited liability companies based in the Netherlands. They were suspected by the Intelligence and Investigation Service of the Ministry of Housing, Spatial Planning and the Environment of having committed the criminal offence of forgery of documents, as well as involvement in the disposal of polluted waste, in contravention of environmental protection legislation, bribery of an official, fraud, and forgery of documents. During the investigations, telephone conversations of certain company employees were intercepted. Because these conversations contained indications of price-fixing and were deemed particularly relevant to the Netherlands Competition Authority (*Nederlandse Mededingingsautoriteit* – “the NMA”), the public prosecutor overseeing the investigation authorized, under section 39f of the Criminal Data Act (*Wet Justitiële en Strafvorderlijke gegevens* – “the WJSG”)⁷, their transmission to the NMA.

As a consequence of that transfer, relying on arts. 8 and 13 ECHR, the applicant companies argued that the Competition Authority’s use of lawfully

governmental organisation or group of individuals claiming to be the victim of a violation by one of the High Contracting Parties of the rights set forth in the Convention or the Protocols thereto. The High Contracting Parties undertake not to hinder in any way the effective exercise of this right».

⁶ ECtHR, *Ships Waste Oil and Others v. Netherlands*, app. no. 2799/16 and 3 others (2025). See R. Mazzocchi, *La causa Ships Waste Oil Collector: un nuovo caso di diritto della concorrenza alla Corte europea dei diritti dell’uomo*, in *Eurojus.it rivista*, 2024, 136 ff.; E. Ward, *Ships Waste Oil Collector B.V. And Others V Netherlands – Article 34 And The Margin Of Appreciation: Companies As Rights-Holders Under Article 8*, in *Strasbourg Observers*, 2025, 1 ff.; S. Margale, *Thin Line Between Protection and Permission: ECtHR’s Endorsement of Evidence Repurposing*, in *Völkerrechtsblog*, 2025, 1 ff.; S. Huijts, *Strasbourg calling – ECHR majority clears competition wiretap case (but mind the dissenters)*, in *The Thicket*, 2025.

⁷ Section 39f, para. 1, WJSG provides: «The Board of Prosecutors-General may ... in so far as it is necessary in view of a compelling general interest (*zwaarwegend algemeen belang*), transmit criminal data to persons or public authorities (*instanties*) for the following purposes:

- (a) preventing and identifying criminal offences;
- (b) maintaining order and safety;
- (c) the enforcement of legislation;
- (d) making decisions under administrative law;
- (e) evaluating the necessity of taking a measure related to legal status or disciplinary

intercepted data irrelevant to the criminal investigations was not foreseeable and that the procedural safeguards were insufficient.

Art. 8, para. 1, ECHR states: «Everyone has the right to respect for his private and family life, his home and his correspondence»⁸; it does not explicitly address the issue of data protection. However, consistent with the evolutionary approach adopted by the European Court of Human Rights, which considers the Convention to be a “living instrument”⁹, the guaran-

action;

(f) providing assistance to victims and others affected by criminal offences; or

(g) the carrying out of a private-law legal act by a person or entity assigned a public task».

It is also important to focus on section 39f, para. 2, WJSG, which states: «The Board of Prosecutors-General may only transmit criminal data to persons or official bodies as referred to in the first paragraph to the extent that those data are, for those persons or official bodies:

(a) necessary in view of a compelling general interest or the determination, exercise or defence of a right in law;

(b) provided in a manner that reasonably prevents the identification of persons other than the individual concerned».

⁸ For an overview of the importance of data protection, see L. Bygrave, *Data Protection Pursuant to the Right to Privacy in Human Rights Treaties*, in *International Journal of Law and Information Technology*, 6(3), 1998, 247 ff.; J. Kokott – C. Sobotta, *The distinction between privacy and data protection in the jurisprudence of the CJEU and the ECtHR*, in *International Data Privacy Law*, 3(4), 2013, 222 ff.; O. Pollicino – M. Bassini, *Art. 8 Protezione dei dati di carattere personale*, in R. Mastroianni – O. Pollicino – S. Allegrezza – F. Pappalardo – O. Razzolini (eds.), *Carta dei diritti fondamentali dell'Unione europea*, Milano, 2017, 134 ff.; A. Terrasi, *Il rapporto tra diritto alla privacy e protezione dei dati personali tra Corte di giustizia dell'Unione europea e Corte europea dei diritti dell'uomo*, in M. Distefano (ed.), *La protezione dei dati personali ed informatici nell'era della sorveglianza globale*, Napoli, 2017, 127 ff.; T. Wisman, *Privacy and Data Protection - Fundamentally Complex*, in *European Data Protection Law Review*, 4(1), 2018, 118 ff.; G. Della Morte, *Big Data e protezione internazionale dei diritti umani*, Napoli, 2018, 174 ff.; M. Paterson – M. McDonagh, *Data protection in an era of big data: The challenges posed by big personal data*, in *Monash University Law Review*, 44(1), 2018, 1 ff.; P. Chhabria, *Right to privacy and data protection*, in *Indian Journal of Law and Legal Research*, 3(2), 2021, 3 ff.; Y. Kovalenko, *The Right to Privacy and Protection of Personal Data: Emerging Trends and Implications for Development in Jurisprudence of European Court of Human Rights*, in *Masaryk University Journal of Law and Technology*, 16(1), 2022, 39 ff.; G.B. Chanal, *International Data Protection Laws*, in *International Journal of Law Management & Humanities*, 6(5), 2023, 2000 ff.; S. Khadzhiradieva – B. Bezverkhniuk – O. Nazarenko – S. Bazyka – T. Dotsenko, *Personal data protection: Between human rights protection and national security*, in *Social Legal Studios*, 7(3), 2024, 247 ff.

⁹ The European Union legal framework acknowledges the protection of personal data as an autonomous fundamental right, explicitly articulated in art. 8 of the Charter of Fundamental Rights of the European Union (CFREU) and distinctly distinguished from the right to respect for “private life” delineated in art. 7 CFREU. This structural separation, however, does not find an equivalent at the international level, where data protection has primarily developed through judicial interpretation. In this particular context, the European Court of Human Rights has progressively expanded the scope of art. 8 ECHR to encompass data protection by adopting an expansive interpretation of the right to “private life”.

In this regard, see C. Pitea – L. Tomasi, *Articolo 8 della Convenzione europea dei diritti dell'uomo*, in S. Bartole – P. De Sena – V. Zagrebelsky (eds.), *Commentario breve alla Convenzione europea per la salvaguardia dei diritti dell'uomo e delle libertà fondamentali*, Padova, 2012, 298 ff.; A. Pisapia, *La tutela multilivello garantita ai dati personali nell'ordinamento europeo*, in *federalismi.it*, 2018, 4 ff.; O. Pollicino, *Judicial Protection of Fundamental Rights on the Internet: A road Towards Digital Constitutionalism?*, Oxford, 2021, 101 ff.; M. Castellaneta, *Articolo 8. Diritto al rispetto della vita privata e familiare*, in R. D'Orazio – G. Finocchiaro – O. Pollicino – G. Resta

tee of private life has been extended to include the protection of privacy and personal data, now recognized as a fundamental right.

However, it should be noted that this right is not absolute; art. 8, para. 2, ECHR permits interference if «it is in accordance with the law and is necessary in a democratic society in the interests of national security, public safety or the economic well-being of the country, for the prevention of disorder or crime, for the protection of health or morals, or for the protection of the rights and freedoms of others»¹⁰.

The Court acknowledged that, despite the differences between natural and legal persons, the minimum guarantees set out in art. 8 of the Convention apply to both. However, given that the data pertained exclusively to commercial relationships, as opposed to aspects of intimate or personal identity, the requisite level of protection was not at its highest.

The Court then proceeded to assess the compliance of the interference with art. 8, para. 2, ECHR. The measure was prescribed by a clear and accessible rule (art. 39f WJSG), pursued a legitimate aim (namely, the economic well-being of the country), and was proportionate, with procedural safeguards including judicial review of prosecutorial authorizations and effective remedies for the affected parties. Furthermore, the Court clarified that the extent of the margin of appreciation is independent of the legal status of the subject, whether natural or legal, and is contingent on the nature of the data, the magnitude of the interference, and the objective pursued. The margin of appreciation is more expansive for data pertaining to commercial activities and more limited for data associated with private or personal domains¹¹.

In addition to this, the applicants contended that they had not been furnished with a suitable remedy as outlined in art. 13 ECHR, but the Court observed that the available civil and administrative remedies were adequate, thereby excluding a combined violation of arts. 8 and 13 of the Convention.

The present analysis seeks to examine the extent to which art. 8 ECHR protects the commercial data of legal persons, with particular attention to the balance between public interests and fundamental rights, and the central role of procedural safeguards in ensuring effective protection.

2. The National Proceedings

The judgments that resulted in the referral to the Grand Chamber concerned *Ships Waste Oil Collector B.V.*, *Janssen de Jong Groep B.V.*, *Janssen de Jong Infra B.V.*, *Janssen de Jong Infrastructuur Nederland B.V.*, *Burando Holding B.V.*, and *Port Invest B.V.*

(eds.), *Codice della privacy e data Protection*, Milano, 2021, 3 ff.

¹⁰ See, on this point, C. Pitea – L. Tomasi, *Articolo 8 della Convenzione europea dei diritti dell'uomo*, cit., 306 ff.; M. Castellaneta, *Articolo 8. Diritto al rispetto della vita privata e familiare*, cit., 7 ff.; C. Jing, *The ECtHR's Suitability Test in National Security Cases: Two Models for Balancing Human Rights and National Security*, in *Leiden Journal of International Law*, 36(2), 2023, 295 ff.

¹¹ See § 162 of the judgment.

The Dutch companies Ships Waste Oil Collector B.V., Burando Holding, and Port Invest are involved in the collection of liquid waste. In 2006, court-authorized wiretaps were conducted in a criminal investigation targeting other operators, and some conversations with employees of the applicants were transmitted to the Netherlands Competition Authority with the prosecutor's authorization. In 2007, similar wiretaps involved Janssen de Jong companies, suspected of bribery, and relevant communications were also transferred to the competent authority.

The applicants were fined based on this evidence and challenged its use, arguing a violation of art. 8 ECHR. The initial ruling by a Dutch court of first instance determined the transfer to be unlawful; however, this decision was subsequently overturned in 2015 on the grounds of public interest and the absence of less intrusive alternatives. The cases were submitted to the European Court of Human Rights on 7 January 2016, and in three judgments of 16 May 2023, a Chamber of the Court found no violation of arts. 8 and 13 ECHR.

After this decision, the applicants requested that the case be referred to the Grand Chamber, which accepted the request.

A Grand Chamber hearing was held on 6 March 2024, and the judgment was delivered on 1 April 2025, confirming that there had been no violation of arts. 8 and 13 ECHR.

3. The Status of Companies as “Victims” under the European Convention on Human Rights

In order to analyze the *Ships Waste Oil* case, it is necessary to examine, first of all, the meaning of the term “everyone”, as used in art. 8, para. 1, ECHR. While the terms “everyone” and “no one” may include legal persons in certain circumstances, they are also used in provisions that are clearly restricted to human beings, such as the right to life (art. 2 ECHR) and the prohibition of torture (art. 3 ECHR). The interpretation of these terms depends on the context: they partly refer exclusively to natural persons and partly to both natural and legal persons¹².

To determine whether legal persons may effectively invoke art. 8 ECHR, it is therefore necessary to assess whether, and under what conditions, commercial companies may be recognized as “victims” entitled to invoke the rights guaranteed by the Convention, under art. 34¹³. The Court has clarified¹⁴ that the criteria for determining whether such status is held cannot be applied in a rigid, mechanical and inflexible manner¹⁵. It follows that any interpretation characterized by “excessive formalism” must be avoided.

¹² L. Dopplinger, *Legal Persons as Bearers of Rights Under the ECHR*, cit., 7.

¹³ See § 146 of the judgment.

¹⁴ See, on this point, M. Ferri, *La decisione sul caso Shortall Et Al. v. Ireland: verso un ulteriore (ed eccessivo) restringimento della nozione di vittima potenziale nel sistema ceduz?*, in *SidiBlog*, 2022.

¹⁵ ECtHR, *Karner v. Austria*, app. no. 40016/98 (2003); ECtHR, *Gorraiz Lizarraga and Others v. Spain*, app. no. 62543/00 (2004), §§ 37-38.

To comprehend the issue in its entirety, it is imperative to undertake a meticulous examination of the article's text, which enables "persons" to file complaints. Although the English text employs the term "person", which is theoretically broad and potentially inclusive of legal persons, the French text refers to "*personne physique*", which denotes natural persons, thus excluding legal entities¹⁶. Consequently, in the event of a discrepancy between language versions, the meaning must be interpreted, in accordance with customary international law as well as the 1969 Vienna Convention on the Law of Treaties¹⁷, in light of the object and purpose of the Convention, and in this case, the French wording must prevail¹⁸.

The principle of effectiveness of the Convention necessitates a broad interpretation of rights; however, this does not necessitate the inclusion of legal persons, whose recognition would contradict the original intent to protect the rights of natural persons. This interpretation is corroborated by a juxtaposition with art. 1, para. 1, Protocol No. 1, which expressly governs instances in which legal persons may be encompassed, and by the historical context in which the notion of human dignity persists as a pivotal element¹⁹.

Despite the restrictive nature of the article's wording, the Court has, on occasion, ruled that a company may initiate legal action if it fulfils the criteria for designation as a "non-governmental organization" (NGO). The distinction between governmental and NGO is predicated on three factors: legal status, activities carried out, and independence from the State. However, the justification for classifying commercial companies as "non-governmental organizations" remains ambiguous and not fully consistent with the Convention's objective, which is to protect human dignity and the rights of "victims", rather than to pursue economic interests. In the preparatory work for the Convention, the initial term "corporate person" was replaced by "non-governmental organization", reflecting a conscious decision to exclude commercial entities.

Nevertheless, the jurisprudence of the Court has evolved in a manner

¹⁶ The French version of art. 34 ECHR provides: «*La Cour peut être saisie d'une requête par toute personne physique, toute organisation non gouvernementale ou tout groupe de particuliers qui se prétend victime d'une violation par l'une des Hautes Parties contractantes des droits reconnus dans la Convention ou ses protocoles. Les Hautes Parties contractantes s'engagent à n'entraver par aucune mesure l'exercice efficace de ce droit*».

¹⁷ See arts. 31 - 33 of the Vienna Convention on the Law of Treaties (23 May 1969). There is an extensive literature regarding the interpretation of treaties; therefore, see the contributions considered most relevant to the analysis: C.B. Kuner, *The interpretation of multilingual Treaties: Comparison of Texts versus the Presumption of similar Meaning*, in *International and comparative Law Quarterly*, 40(4), 1991, 954 ff.; G. Strozzi, *Il diritto dei trattati*, Torino, 1999, 61 ff.; O. Dörr, *Article 33*, in O. Dörr, K. Schmalenbach (eds.), *Vienna Convention on the Law of Treaties – A Commentary*, Heidelberg, 2012, 596 ff.; E. Cannizzaro, *Diritto Internazionale*, Torino, 2012, 176 ff.; N.R. Doshi, *The Law of Treaties with Special Focus on Evolutionary Interpretation of the Treaties*, in *International Journal of Law Management & Humanities*, 4(3), 2021, 2596 ff.

¹⁸ ECtHR, *Wemhoff v. Germany*, app. no. 2122/64 (1968). In this regard, see L. Pasquali, *On the Interpretation of Multilingual Treaties: The Need to Resort to Specific Criteria*, in *Osservatorio sulle fonti*, 3, 2017, 37.

¹⁹ E. Ward, *Ships Waste Oil Collector B.V. And Others V Netherlands – Article 34 And The Margin Of Appreciation: Companies As Rights-Holders Under Article 8*, cit., 5.

that is increasingly inclusive. Since the 1970s, the interpretation of the Convention has undergone a transformation, moving from limited protection for natural persons to protection that also extends to legal persons. Two specific cases can be cited as exemplars: *Sunday Times v. United Kingdom*²⁰, concerning freedom of expression, and *Sociétés Colas Est v. France*²¹, in which art. 8 ECHR was applied to the “domicile” of the company. The jurisprudence on this subject was further consolidated in the *Vinci Construction v. France* judgment²², concerning the lawfulness of the mass seizure of company emails in the context of an antitrust investigation.

This development is founded on the principle articulated in art. 1, para. 1, ECHR, which provides that States are obligated to ensure the rights and freedoms of “everyone” within their jurisdiction, without confining the scope of protection exclusively to natural persons. However, in the case under analysis, as Judge Serghides’ dissenting opinion pointed out, a regulatory gap remains²³. The Court has not clarified the legal basis that would allow companies to lodge individual applications under art. 34 ECHR. The applicants were commercial companies, which cannot easily be subsumed under any of the three “windows” of the provision. While the “NGO” window is the least problematic, as companies established as non-profit organizations can apply under this category, the “person” window, bound by the French term “*personne physique*”, and the “group of individuals” window exclude commercial companies from “victim” status. Furthermore, the dissenting judge issues a *caveat* that this very principle may, in the future, hinder the Court’s ability to extend its protection to forms of artificial intelligence that exhibit functional capabilities analogous to those of natural persons. If States elect to confer rights upon artificial intelligence, this should be achieved «through the implementation of a new Protocol»²⁴. The purpose of this Protocol would be to clarify the legal foundation for entities other than natural persons, in accordance

²⁰ ECtHR, *The Sunday Times v. the United Kingdom*, app. no. 6538/74 (1979).

²¹ ECtHR, *Société Colas Est and Others v. France*, app. no. 37971/97 (2002).

²² ECtHR, *Vinci Construction and GTM Génie Civil et Services v. France*, app. nos. 63629/10 and 60567/10 (2015).

²³ See § 14 of the dissenting opinion of Judge Serghides.

²⁴ In this matter, §§ 48 - 49 of the dissenting opinion of Judge Serghides, who considered: «If, one day, member States decide to provide some kind of protection for artificial intelligence, it would be better done through a new Protocol. A new Protocol would also offer the opportunity to firmly and clearly set out a legal basis for the claims of entities other than natural persons, including companies. However, one can only wonder how such developments in the future, if they came to be, could be reconciled with the Convention’s primary aim for the effective protection of human rights and the principle of human dignity! If, in the future, artificial intelligence has rights similar or adjacent to human rights safeguarded by the Convention, as granted by a new Protocol to the Convention, how could one be certain that it would itself respect human rights, the rule of law and the principle of democracy. Though it may be conceivable that autonomous artificial intelligence may in the future benefit from the protection of rights similar or adjacent to human rights contained in the Convention, at the same time, it is crucial that this goes hand in hand with the capacity and obligation of such artificial intelligence to respect human rights, the rule of law and the principle of democracy. However, I cannot continue further this interesting, though theoretical and philosophical discussion, since this is not the issue here».

with human rights, the rule of law, and democratic principles.

This contrast between theoretical and practical approaches was underlined in the *Ships Waste Oil* case, which brought the issue of protecting companies' rights to the fore. Against this background, the companies claimed "victim" status based on an alleged interference under art. 8, para. 2, ECHR.

As will be analyzed in the following sections, the Court recognized that the interception of data and its transmission to the competition authority constituted an interference with the right to respect for correspondence, but it was considered proportionate and justified, since the data concerned exclusively commercial information on the economic activities of undertakings and not aspects of the private lives of individuals²⁵.

This pragmatic approach can engender paradoxical effects: an employee whose business communications are intercepted has reduced protection if the transmission concerns commercial data, but full protection if it concerns private data. The result is a system of "variable geometry" safeguards, which poses challenges to the principle of the universality of rights²⁶. The case thus highlights an important problem: are corporate rights full-fledged human rights or functional guarantees, recognized to strengthen the rule of law but not equivalent in intensity and content to those of natural persons?

The response to this question carries with it significant practical ramifications. The delineation between personal and commercial data is not solely indicative of the sensitivity of the information; it also reflects the legal status of the subject. Individuals benefit from robust protections against intrusions into their most private sphere²⁷, while legal entities enjoy functional guarantees sufficient to protect the integrity of commercial operations and the rule of law. However, these guarantees are not as weighty as those afforded to fundamental rights of natural persons. Consequently, the system necessitates meticulous balancing to reconcile the protection of privacy, the demands of security, and the recognition of corporate interests, without undermining the coherence and universality of fundamental rights.

4. The Margin of Appreciation and the Protection

²⁵ R. Mazzocchi, *La causa Ships Waste Oil Collector: un nuovo caso di diritto della concorrenza alla Corte europea dei diritti dell'uomo*, cit., 138 ff., 144.

²⁶ There is an extensive literature on the subject of "variable geometry"; therefore, see the contributions considered most relevant to the analysis: Y. Arai-Takahashi, *The margin of appreciation doctrine: a theoretical analysis of Strasbourg's variable geometry*, in A. Føllesdal – B. Peters – G. Ulfstein (eds.), *Constituting Europe: The European Court of Human Rights in a National, European and Global Context. Studies on Human Rights Conventions*, Cambridge, 2013, 62 ff.; B. ÇALI, *Coping with Crisis: Whither the Variable Geometry in the Jurisprudence of the European Court of Human Rights*, in *Wisconsin International Law Journal*, 35(2), 2018, 269 ff.; O. Bajgar – J. Horenovsky, *Negative Human Rights as a Basis for Long-term AI Safety and Regulation*, in *Journal of Artificial Intelligence* 76(1), 2023, 1043 ff.

²⁷ R. Pisillo Mazzeschi, *Diritto internazionale dei diritti umani. Teoria e prassi*, Torino, 2020, 314 ff.

of Strong Public Interests

Once the issue of “victim” status for legal persons, whose personal data were processed has been examined, it is important to consider the margin of appreciation²⁸ granted to States regarding the right to privacy, which includes the protection of personal data. It should first be clarified that both the case law of the European Court of Human Rights and legal doctrine recognize that the margin of appreciation applies to all rights contained in the ECHR.

According to the Court’s established case law, while Member States bear the primary responsibility for securing Convention rights, the ECtHR retains a supervisory role in assessing the legality, necessity, and proportionality of any interferences. Within this framework, States may be granted a margin of appreciation whose breadth is not fixed but rather varies depending on factors such as the existence of a European consensus, the nature of the right in question, the sensitivity of the interests involved, and the seriousness of the intervention²⁹.

In *Ships Waste Oil*, the intercepted data had been lawfully obtained through methods compatible with art. 8; indeed, the criminal data, as mentioned in the first paragraph, had been lawfully transmitted to the Netherlands Competition Authority, under section 39f, para. 1, lit. c), WJSG, the WJSG Instructions, and section 6, para. 1, Dutch Competition Act³⁰.

According to art. 8, para. 2, ECHR, interference by public authorities in the exercise of the right to private life is only permissible «if it is in accordance with the law» and constitutes, as already said, a measure that «is necessary in a democratic society in the interests of national security, public safety or the economic well-being of the country, for the prevention of disorder or crime, for the protection of health or morals, or for the

²⁸ Regarding the margin of appreciation, see J. Sweeney, *A «margin of appreciation» in the internal market: lessons from the European Court of Human Rights*, in *Legal Issues of Economic Integration*, 34(1), 2007, 27 ff.; N. Shuibhne, *Margin of appreciation: national values, fundamental rights and EC free movement law*, in *European Law Review*, 32(2), 2009, 230 ff.; I. Anrò, *Il margine di apprezzamento nella giurisprudenza della Corte di giustizia dell’Unione europea e della Corte europea dei Diritti dell’uomo*, in A. Oddenino – E. Ruozzi – A. Viterbo – F. Costamagna – L. Mola – L. Poli (eds.), *La funzione giurisdizionale nell’ordinamento internazionale e nell’ordinamento comunitario: atti dell’Incontro di studio tra i giovani cultori delle materie internazionalistiche*, Torino, 2009, 7 ff.; O.M. Arnardóttir, *Res Interpretata, Erga Omnes Effect and the Role of the Margin of Appreciation in Giving Domestic Effect to the Judgments of the European Court of Human Rights*, in *European Journal of International Law*, 28(3), 2021, 820 ff.; C.A. Chagas, *Balancing competences and the margin of appreciation: structuring deference at the ECtHR*, in *Vienna Journal on International Constitutional Law ICL Journal*, 16(1), 2022, 19; M.Y. Abdulmumin, *An overview of the doctrine of margin of appreciation in international human rights jurisprudence*, in *International Review of Law and Jurisprudence*, 4(2), 2022, 168 ff.; A. Todorovic, *Whose margin of appreciation? analyzing its significance in european human rights jurisprudence*, in *Kultura Polisa*, 20(2), 2023, 130 ff.

²⁹ P. Friedl, *Reasonable Expectations of Privacy*, in *Issues in Privacy and Data Protection*, Berlin, 2025, 249 ff., 252.

³⁰ Art. 6, para. 1, Dutch Competition Act states: «Agreements between undertakings, decisions by associations of undertakings and concerted practices by undertakings which have as their object or effect the prevention, restriction or distortion of competition within the Dutch market, or a part thereof, are prohibited».

protection of the rights and freedoms of others»³¹.

The Court clarified that «the wording “in accordance with the law” requires the impugned measure both to have some basis in domestic law and to be compatible with the rule of law, which is expressly mentioned in the Preamble to the Convention and inherent in the object and purpose of art. 8 ECHR. The law must therefore meet quality requirements: it must be accessible to the person concerned and foreseeable as to its effects»³².

As regards the lawfulness of the transmission, the Chamber found that it had a legal basis under Dutch law. It further found that the domestic law met the requirement of “foreseeability”, since domestic law gave a clear indication as to the circumstances and the conditions under which prosecutors were empowered to transmit data to third parties under section 39f WJSG³³.

In the Court’s settled case-law, «the interference will be considered “necessary in a democratic society” for a legitimate aim if it answers a “pressing social need” and if it is proportionate to the legitimate aim pursued and if the reasons adduced by the national authorities to justify it are “relevant and sufficient”»³⁴.

The European Court of Human Rights acknowledged the legitimacy of national authorities’ intervention in the suppression of anti-competitive practices, emphasizing that the effective application of competition law constitutes a strong public interest. For this reason, the Court underscored that collaboration between competition authorities and the enforcement of anti-competitive practices is instrumental not only to the proper functioning of the market, but also to overall economic well-being.

Based on this consideration, the Court found no violation of the Convention in the *Ships Waste Oil* case, as the transmission of data was deemed justified considering the legitimate objectives outlined in art. 8, para. 2, ECHR.

In light of the reasoning presented thus far, it should be noted that the Court has also assessed the scope of the margin of appreciation with reference to the criterion of necessity.

It is very common for the margin of appreciation to be assessed also considering “the necessity test”³⁵. The ECtHR, in fact, in *Dudgeon v. the United Kingdom*³⁶, for example, introduced the approach whereby the so-called strict necessity test is applied, requiring that the interference with rights be strictly proportionate to the objective pursued.

In *Ships Waste Oil*, the Court recognized that the margin would tend to

³¹ See § 152 of the judgment.

³² ECtHR, *Roman v. Zakharov v. Russia*, app. no. 47143/06 (2015), § 228.

³³ See § 109 of the judgment.

³⁴ ECtHR, *Coster v. The United Kingdom*, app. no. 24876/94 (2001); *S. and Marper v. the United Kingdom*, app. nos. 30562/04 and 30566/04 (2008), § 101.

³⁵ Regarding the importance of the necessity test, see V.A. Da Silva, *Standing in the Shadows of Balancing: Proportionality and the Necessity Test*, in *International Journal of Constitutional Law*, 20(5), 2022, 1746 ff.; F. Concarella, *La riduzione progressiva del margine di apprezzamento statale nella giurisprudenza della Corte europea dei diritti*, in *Ordine internazionale e diritti umani*, 4, 2025, 948 ff.

³⁶ ECtHR, *Dudgeon v. United Kingdom*, app. no. 7525/76 (1981).

be narrower where the right at stake is crucial to the individual's effective enjoyment of intimate or key rights.

Applying this reasoning, it can easily be argued that sensitive personal data requires a higher level of protection³⁷ because, by its nature, such data tends to reveal particularly intimate aspects of an individual's life and its use may lead to discriminatory effects, damage to individual dignity, or serious and irreversible intrusion with privacy. Technological advances and the growing ability to correlate large amounts of information further exacerbate these risks, allowing sensitive information to be extracted even from non-sensitive data³⁸. For these reasons, the European data protection framework recognizes the need for a strengthened protection regime, distinct from that applicable to ordinary personal data, as an essential requirement for ensuring effective protection of fundamental rights³⁹.

In the case at hand, the Court states that «the breadth of the margin of appreciation should depend in each case on the content and nature of the data in question rather than on the applicants' physical or legal nature or their status»⁴⁰. It will be wider in respect of the collection and processing of business-related data of both companies and individuals than in the case of the collection and processing of data concerning an individual's intimate sphere or a particularly important facet of an individual's existence or identity. The breadth of the margin of appreciation will also depend on the gravity of the interference and the object pursued by it.

Nonetheless, it is still possible that cases with very important rights will end up with a wide margin, depending on how the other factors influence the weight of the competences at stake⁴¹. Along these lines, the Court must duly justify its decision based «on all the circumstances of the case»⁴²,

³⁷ ECtHR, *Glukhin v. Russia*, app. no. 11519/20 (2023). In this judgment, the Court recognizes the particularly sensitive nature of data relating to participation in peaceful protests. In particular, see § 76, in which the Court held: «Personal data revealing political opinions, such as information about participation in peaceful protests, fall within the special categories of sensitive data attracting a heightened level of protection».

³⁸ Regarding the distinction between different types of data, see P. Quinn – G. Malgieri, *The Difficulty of Defining Sensitive Data-The Concept of Sensitive Data in the EU Data Protection Framework*, in *German Law Journal*, 22(8), 2021, 1589 ff.; M.L. Montagnani – M. Verstraete, *What makes data personal?*, in *UC Davis Law Review*, 56(3), 2023, 1165 ff.

³⁹ See, on this point, C. Pitea – L. Tomasi, *Articolo 8 della Convenzione europea dei diritti dell'uomo*, in S. Bartole – P. De Sena – V. Zagrebelsky (eds.), *Commentario breve alla Convenzione europea per la salvaguardia dei diritti dell'uomo e delle libertà fondamentali*, cit., 304; V. Zagrebelsky – R. Chenal – L. Tomasi, *Manuale dei diritti fondamentali in Europa*, Bologna, 2019, 295 ff.

⁴⁰ See § 164 of the judgment.

⁴¹ C.A. Chagas, *Balancing competences and the margin of appreciation: structuring deference at the ECtHR*, cit., 5.

⁴² In this regard, see ECtHR, *Big Brother Watch and Others*, app. nos. 58170/13, 62322/14, 24960/15 (2021). In this judgment the Court affirmed the importance of the evaluation of the circumstances. Specifically, at § 339 it is stated: «In view of the risk that a system of secret surveillance set up to protect national security (and other essential national interests) may undermine or even destroy the proper functioning of democratic processes under the cloak of defending them, the Court must be satisfied that there are adequate and effective guarantees against abuse. The assessment depends on all the circumstances of the case, such as the nature, scope and duration of the possible measures, the grounds

not just those it chooses to deem most important.

This logic, reflecting the margin of appreciation accorded to States, grounded in a case-by-case assessment, also emerges in the case-law of other European Courts. Indeed, despite originating from different regulatory contexts, the case law of the European Court of Human Rights and that of the Court of Justice of the European Union (CJEU) show certain similarity in reasoning directed at giving States some margin of appreciation in the implementation of guaranteed rights⁴³. In that regard, both Courts recognize a margin of autonomous assessment for national authorities, and the breadth of this margin varies depending on the nature of the interference, the type of data concerned, and the objectives pursued. Nevertheless, the CJEU has pointed out that Member States possess circumscribed discretion in curtailing the confidentiality of electronic communications. Indeed, in its judgment in *Lietuva vs. Generalinis prokuroras*⁴⁴, the Court clarified that data retained by electronic communications service providers for the purpose of combating serious crime cannot be utilized for less significant purposes, such as administrative investigations into irregularities by a public prosecutor.

In *Ships Waste Oil*, the dissenting judges based their reasoning precisely on this principle, noting that it did not give sufficient weight to the purpose of data transmission when assessing the State's margin of appreciation and the Court's proportionality review⁴⁵.

required for ordering them, the authorities competent to authorise, carry out and supervise them, and the kind of remedy provided by the national law».

⁴³ On this point, see I. Anrò, *Il margine di apprezzamento nella giurisprudenza della Corte di giustizia dell'Unione europea e della Corte europea dei Diritti dell'uomo*, cit., 8; L.S. Rossi, *I rapporti fra la Carta dei diritti fondamentali e la CEDU nella giurisprudenza delle rispettive Corti*, in *I Post di AISDUE*, 2, 2020, 48 ff.; E. Psychogiopoulou, *Judicial dialogue and digitalization: CJEU engagement with ECtHR case law and fundamental rights standards in the EU*, in *Journal of Intellectual Property, Information Technology and Electronic Commerce Law*, 13(2), 2022, 151 ff.

⁴⁴ CJEU, C-162/22, *A. G. v. Lietuvos Respublikos generalinė prokuratura* (2023). In this ruling, the Court affirms that data retained for the purpose of combating serious crimes cannot be used for less serious purposes. In this matter, see § 41, in which the Court held: «Art. 15(1) of Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002, concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications), as amended by Directive 2009/136/EC of the European Parliament and of the Council of 25 November 2009, read in the light of Arts. 7, 8 and 11 and Art. 52(1) of the Charter of Fundamental Rights of the European Union, must be interpreted as precluding the use, in connection with investigations into corruption-related misconduct in office, of personal data relating to electronic communications which have been retained, pursuant to a legislative measure adopted under that provision, by providers of electronic communications services and which have subsequently been made available, pursuant to that measure, to the competent authorities for the purpose of combating serious crime».

⁴⁵ In this matter, see § 3 of the dissenting opinion of Judge Arnardóttir, joined by Judges Serghides and Šimáčková, which emphasized: «The Grand Chamber should have stated more clearly that the transmission of intercept criminal data for purposes unrelated to the original purpose of the interception can only be justified in pursuit of weighty interests. We regret that the majority have opted for an approach that may result in a considerably lower level of protection than that provided under the ePrivacy Directive (Directive 2002/58/EC, as amended by Directives 2006/24/EC and 2009/136/EC) of the European Parliament and of the Council of 12 July 2002 concerning the processing

In contrast to this approach, the Grand Chamber reaffirmed the legitimacy of intervention by national authorities, recognizing that the effective application of competition law is a strong public interest.

A similar balancing logic can be found, albeit with divergent outcomes, in the *Bărbulescu* case⁴⁶.

Mr. Bogdan Mihai Bărbulescu was a worker who complained of the violation of art. 8 ECHR, because he was fired due to the discovery, from his boss, that he had utilized the company's Yahoo Messenger account for personal reasons during business hours. This conduct was in direct contravention of the company's established regulations, which explicitly prohibited the utilization of company equipment for non-work-related purposes⁴⁷.

Consequently, Mr. Bărbulescu initiated legal proceedings in the national courts to obtain a declaration that his dismissal was unlawful. After the exhaustion of all domestic remedies, the case was brought before the European Court of Human Rights, because he sustained that the Romanian State and its courts had failed to protect the right to respect for his private life and correspondence.

In the initial *Bărbulescu* ruling, the European Court of Human Rights, citing its own case law⁴⁸, reaffirmed that communications and emails sent from the workplace are encompassed by the concepts of "private life" and "correspondence" protected by art. 8 of the Convention.

In the case at hand, the Fourth Chamber determined that the employer's surveillance of its employees' correspondence was permissible, without delving into the specific existence of domestic legislation that would justify such restrictions.

The Grand Chamber, however, subsequently overturned this approach⁴⁹, criticizing the inadequacy of the safeguards put in place at the domestic level and underlining the need for national authorities to exercise stricter

of personal data and the protection of privacy in the electronic communications sector (OJ 2002 L 201). We note in this regard that the ePrivacy Directive provides protection for the legitimate interests of legal persons (see Art. 1(2) of the Directive). As interpreted by the CJEU, it further provides that intercept data cannot in principle be transmitted for use in pursuit of purposes of lesser importance than those which justified the interception».

⁴⁶ ECtHR, *Bărbulescu v. Romania*, app. no. 61496/08 (2017).

⁴⁷ For further informations of the importance of *Bărbulescu* case, see J. Eichenhofer, *Internet privacy at work the ECtHR Barbulescu judgment*, in *European Data Protection Law Review*, 2(2), 2016, 268; E.C. Grigorie, *Protecting Privacy and the Change of ECHR Jurisdiction in Barbulescu against Romania*, in *Revista de Stiinte Juridice*, 2, 2020, 230 ff.

⁴⁸ ECtHR, *Niemietz v. Germany*, app. no. 13710/88 (1992); ECtHR *Copland v. the United Kingdom*, app. no. 62617/00 (2007).

In these judgments the Court determined that respect for private life must also be protected in the context of employment relationships.

⁴⁹ See § 140 of the judgment, which indicates the errors made by the Court of Appeal, including the failure to assess the notification of the wiretapping of communications, as well as the absence of any investigation into the nature and extent of these. The Court also noted: «They failed to determine, firstly, the specific reasons justifying the introduction of the monitoring measures; secondly, whether the employer could have used measures entailing less intrusion into the applicants private life and correspondence; and thirdly, whether the communications might have been accessed without his knowledge».

control over compliance with art. 8 ECHR. In doing so, the Court highlighted that the recognition of a margin of appreciation does not exempt States from their obligation to ensure an adequate level of protection of fundamental rights, reaffirming its ultimate supervisory role.

Therefore, in both *Ships Waste Oil* and *Bărbulescu*, the State's margin of appreciation is closely linked to the correct balancing of the interests at stake⁵⁰. This may lead the ECtHR to uphold national decisions that appear proportionate and supported by one of the justifications set out in art. 8, para. 2, ECHR. However, it may give rise to a finding of a violation of the Convention when that balance is inadequate or excessively biased in favour of one party.

5. When Interference Becomes a Violation: a Comparison of Two Approaches

Following this discussion on corporate rights and on the margin of appreciation, the European Court of Human Rights' case law on wiretapping and searches illustrates that the protection afforded by art. 8 ECHR relies more on the quality of the legal framework and procedural safeguards than on the intrusive nature of the measures themselves.

In *Sociétés Colas Est and Others v. France*, for example, the Court criticized the French legal system for ordering company searches without prior judicial supervision or clear limits, resulting in them being disproportionate. Therefore, interference can only be tolerated if it is part of a predictable, accessible and proportionate regulatory system.

The judgment in *Ships Waste Oil* is consistent with this perspective and provides a useful point of comparison. In this case, despite the transfer of wiretap recordings from the judicial authorities to the NMA, the Court ruled that there had been no violation of art. 8 ECHR. This was justified by the existence of a regulatory framework (the WJSG law), which provided for the transfer of intercepted data to other competent authorities. While the NMA was not expressly mentioned, the Court deemed the transmission to be 'sufficiently foreseeable' and supported by procedural safeguards to prevent abuse⁵¹.

Furthermore, the cooperation between the public prosecutor's office and the NMA was limited to selecting data relevant to matters of general interest, with no indiscriminate access to additional information. The Court therefore concluded that prior judicial authorization was unnecessary, as the system was capable of ensuring an adequate balance.

However, the ruling was not unanimous. The case was referred to the Grand Chamber, where three judges expressed dissenting opinions of considerable importance. According to these judges, the original author-

⁵⁰ See B. Van Der Sloot, *Privacy and Personality Right: Why the ECtHR's Focus on Ulterior Interests Might Prove Indispensable in the Age of "Big Data"*, in *Utrecht Journal of International and European Law*, 31(80), 2015, 25 ff.; Y. Kovalenko, *The Right to Privacy and Protection of Personal Data: Emerging Trends and Implications for Development in Jurisprudence of European Court of Human Rights*, cit., 47.

⁵¹ See §§ 117-124 of the judgment.

ization could not automatically justify the use of intercepted data for administrative purposes unrelated to the original criminal investigation. They underscored the need for specific safeguards for data transfers between authorities, invoking the principles of proportionality and minimization: only information that was strictly necessary for the secondary investigation should have been transmitted.

From an EU legal perspective, they also pointed out that the transfer of data must comply with the principles of purpose limitation and legitimacy, as set out in the Directive on privacy and electronic communications⁵². The Court of Justice of the European Union has already clarified that declassifying data collected in criminal proceedings for disciplinary or administrative purposes is likely to be incompatible with European law. Consequently, the dissenting judges concluded that the prosecutor's office should have provided more compelling reasons to justify the seriousness of the non-criminal violations being prosecuted and should have ensured effective judicial review of the transmission.

The contrast between the majority and dissenting opinions in *Ships Waste Oil* highlights the sensitivity of the issue. The Court appears willing to permit the use of data obtained during criminal investigations for other purposes, provided there is a robust regulatory framework in place and adequate safeguards are in force. However, the minority argues for more stringent limitations, particularly in view of EU legislation.

This makes the ruling particularly significant for comparison with other legal systems.

In this regard, it is relevant to compare the *Ships Waste Oil* case with the judgment handed down against Italy on 23 May 2024⁵³. On that occasion, the ECtHR ruled that intercepting communications in criminal proceedings in which the applicant was not a defendant violated art. 8 ECHR. The critical issue was not the use of wiretaps *per se*, but rather the absence of an effective legal remedy that would allow someone not involved in the proceedings to challenge the measure. Although the applicant became aware of the wiretaps through the search warrant, he had no legal means of seeking a judicial review. In line with the judgment in *Roman Zakharov v. Russia*⁵⁴, the Court recalled that the possibility of challenging a surveillance measure is an essential safeguard against potential abuse. It was precisely this absence that led Strasbourg to censure Italian legislation as deficient in terms of the "quality of the law".

The comparison between the two cases is illuminating. In the Netherlands, the Court highlighted the clarity of the legal basis and the presence of procedural safeguards, ruling out a violation of art. 8 despite the interference. In Italy, however, the absence of judicial remedies rendered the intervention disproportionate and incompatible with a democratic society.

⁵² Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications).

⁵³ ECtHR, *Contrada v. Italy*, app. no. 2507/19 (2024).

⁵⁴ ECtHR, *Zubkov and Others v. Russia*, app. nos. 29431/05 (2018), § 94.

6. The Importance of Art. 39 of the Dutch Judicial and Criminal Records Data Act (WJSG)

The judgment is relevant because the Court underscored that the measure was based on a solid national legal framework, namely art. 39f, WJSG, which defines the conditions under which the public prosecutor may transfer data to other public authorities for purposes of general interest, such as the protection of competition.

These conditions include significant public interest, necessity for a specific purpose, and impossibility of obtaining the data through less intrusive means. The Court underlined that, while the transfer was classified as a factual act rather than a formal decision under Dutch administrative law, procedural safeguards were available afterwards through standard legal remedies and administrative proceedings.

This framework was considered compatible with art. 8 ECHR, as it ensures judicial oversight, predictability, and protection against arbitrariness. In contrast, in the case of *Brazzi v. Italy*⁵⁵, the Court noted that a measure with a formal legal basis could still violate art. 8 if effective judicial review was lacking, whether prior or subsequent. In this instance, the absence of adequate safeguards meant that the interference with the applicant's home was unlawful, thereby emphasizing the vital function of judicial oversight in ensuring compliance with the legality requirement under art. 8, para. 2, ECHR⁵⁶. This general principle sheds light on the critical issues in the *Ships Waste Oil* case, where the data transfer, despite being based on solid legal grounds, is classified as a factual act.

The Chamber recognized that transmitting intercepted data to the NMA, as carried out by the employees of the applicant companies, constituted an interference with the companies' rights under art. 8 of the Convention, but this was lawful and proportionate. The transfer of data was based on Dutch law, was foreseeable, and the legal and judicial safeguards - including the possibility of *ex post* review by civil and administrative courts - ensured that the intervention was compatible with a democratic society and pursued a legitimate public interest: the protection of economic well-being.

However, the dissenting opinions criticized this analysis, arguing that authorizations for the transfer of data lacked written reasoning on necessity and proportionality⁵⁷. Since Dutch law treated the transfer of data as a factual act rather than a formal decision, *ex post* judicial review could not substitute for prior assessment, rendering safeguards against arbitrariness and abuse ineffective. According to the dissenting judges, this deficiency should have resulted in a finding of violation, as the national system failed to restrict data transfers to what was strictly necessary in a democratic

⁵⁵ ECtHR, *Brazzi v. Italy*, app. no. 57278/11 (2018).

⁵⁶ A. Viotto, *Il "diritto al rispetto della vita privata e familiare" nell'ambito delle indagini tributarie, nel quadro della giurisprudenza della Corte Europea dei Diritti dell'Uomo*, in *Rivista Trimestrale di diritto tributario*, 1, 2019, 155 ff.

⁵⁷ See § 10 of the dissenting opinion of Judge Arnardóttir, joined by Judges Serghides and Šimáčková.

society. While the Chamber considered the Dutch framework sufficient to balance public and private interests, the dissenters stressed that the absence of written reasoning and broad prosecutorial discretion undermined protections.

In *Ships Waste Oil*, this classification has immediate practical consequences for the companies involved, potentially affecting their legal, economic, and reputational interests without prior justification, adversarial procedure, or immediate judicial review. The case illustrates that formal legality and general safeguards do not prevent immediate harm to rights. The legal classification of administrative and investigative acts is therefore not a mere technicality; it profoundly shapes the level of protection afforded to fundamental rights, demonstrating the importance of reconciling efficient public action with effective safeguards for individual rights.

7. Concluding Remarks

Considering the analysis conducted, the judgment in *Ships Waste Oil* raises important questions regarding the application and interpretation of art. 8 ECHR.

The applicant companies were found guilty of multiple criminal offenses, including document forgery, illegal disposal of polluted waste, bribery, fraud, and price fixing⁵⁸, a practice that undermines market competition and harms consumers. These actions violated both domestic law, under Section 6, para. 1, Dutch Competition Act, and European law, particularly art. 101 TFEU⁵⁹, which prohibits agreements that distort competition within the internal market.

In the case at hand, the Court emphasized the strong public interest involved in the effective enforcement of competition law because it is important not only to safeguard the functioning and fairness of markets, but also to protect the overall economic well-being of society. Given the potential negative impact of cartels on market competition and the inherent difficulties in detecting and investigating them, it is essential that competition authorities and other law enforcement agencies cooperate in their efforts to uncover and punish such anti-competitive practices.

⁵⁸ G. Day, *Is the Problem with Antitrust Law or Antitrust Enforcement?*, in *Journal of Law and Political Economy*, 5(1), 2025, 120 ff.

⁵⁹ Art. 101, para. 1, TFEU provides: «The following shall be prohibited as incompatible with the internal market: all agreements between undertakings, decisions by associations of undertakings and concerted practices which may affect trade between Member States and which have as their object or effect the prevention, restriction or distortion of competition within the internal market, and in particular those which:

- (a) directly or indirectly fix purchase or selling prices or any other trading conditions;
- (b) limit or control production, markets, technical development, or investment;
- (c) share markets or sources of supply;
- (d) apply dissimilar conditions to equivalent transactions with other trading parties, thereby placing them at a competitive disadvantage;
- (e) make the conclusion of contracts subject to acceptance by the other parties of supplementary obligations which, by their nature or according to commercial usage, have no connection with the subject of such contracts».

However, the Court's decision carries a significant risk of weakening the fundamental principle of purpose limitation, which is a cornerstone of data protection in Europe⁶⁰. This principle guarantees the data subject that the use and purpose of processing personal data are the ones that were specifically previously determined and informed by the controller to the data subject, before starting with the processing of data. It is an important data protection principle, created to provide transparency to data processing, to assess the legitimacy of data processing, and to prevent the abuses of the controller's powers.

By allowing, in *Ships Waste Oil*, the reuse of data lawfully collected in the context of criminal investigations for separate administrative purposes, without new judicial authorization or an individual notification requirement, the Grand Chamber reduces predictability and control for data subjects. Although the Court does not explicitly address EU data protection instruments, its tolerance of functional data reuse could gradually erode this principle.

This risk is also confirmed by the dissenting opinions, which illustrated how the Court's focus on the effective application of competition rules has led to tolerance of a reuse of intercepted data that lacks a clear definition in regulatory and procedural terms. National law should set a minimum level of gravity regarding cases that may justify the further transfer of data and information collected during criminal investigations. A similar approach is necessary to avoid circumvention of strict safeguards on means of gathering evidence and to prevent "fishing expeditions"⁶¹ by law enforcement agencies and public authorities in general. It would also have been preferable for the prosecutor's instructions to have provided more detailed and specific guidance on the level of seriousness of non-criminal

⁶⁰ The principle of limitation of purpose is expressly set out in arts. 5, para.1, lit. b), and 6, para. 4, Regulation (EU) 2016/679.

See E. Brouwer, *Legality and Data Protection Law: The Forgotten Purpose of Purpose Limitation*, in L.F.M. Besselink – F. Pennings – S. Prechal (eds.), *The Eclipse of the Legality Principle in the European Union*, Kluwer, 2011, 273 ff.; C. Jasserand, *Subsequent use of GDPR data for law enforcement purpose: the forgotten principle purpose limitation?*, in *European Data Protection Law Review* 4(2), 2018, 152 ff.; L. Jasmontaite – I. Kamara – G. Zanfiri-Fortuna – S. Leucci, *Data Protection by Design and by Default: Framing Guiding Principles into Legal Obligations in the GDPR*, in *European Data Protection Law Review* 4(2), 2018, 168 ff.; I. Hahn, *Purpose Limitation in the Time of Data Power: Is There a Way Forward?*, in *European Data Protection Law Review*, 7(1), 2021, 31 ff.; D. Baldini, *The impact of the right to personal data protection on the design of the european digital identity wallet*, in *Italian Journal of Public Law*, 16(1), 2024, 297 ff.

⁶¹ E. G. Thornburg, *Just Say "No" to Fishing: The Lure of the Metaphor*, in *University of Michigan Journal of Law Reform*, 40(1), 2006, 35, notes that «the term "fishing expedition" was coined in *Hickman v. Taylor* 329 U.S. 495 (1947), to describe discovery requests lacking a sufficient factual basis and aimed at uncovering previously unknown or unpleaded evidence».

See also M. Michalek-Gervais, *Fishing Expeditions and Subsequent Electronic Searches in the Light of the Principle of Proportionality of Inspections in Competition Law Cases in Europe*, in *ASCOLA Working Paper*, Warsaw, 7(10), 2014, 3 ff.; L. Keunen, *Big data tax audits: the conceptualisation of fishing expeditions*, in *International Review of Law, Computers & Technology*, 37(2), 2023, 166 ff.; R. Mazzocchi, *La causa Ships Waste Oil Collector: un nuovo caso di diritto della concorrenza alla Corte europea dei diritti dell'uomo*, cit., 142 ff.; L. Arnold – D. Bhatnagar – C. Shen – E. Bershtskiy, *Use of Facial Recognition Technologies for Law Enforcement: Comparative Analysis*, in *Oregon Review of International Law*, 26(175), 2025, 178 ff.

offenses that would allow the transmission of wiretaps, and for the prosecutor's office itself to have provided at least a "properly reasoned decision" regarding the interference with the right referred to in art. 8 ECHR. This is a lack of reasoning that does not allow, among other things, for adequate *ex post* judicial review⁶².

Therefore, accepting the reuse of data in the absence of clear limits, reasoned authorizations, and preventive safeguards means exposing the principle of purpose limitation to progressive erosion, with effects that go beyond the individual case and affect the future interpretation of art. 8 ECHR.

Keywords

data protection - strong public interest - criminal vs. administrative enforcement - margin of appreciation - fundamental rights and legal persons

⁶² S. Margale, *Thin Line Between Protection and Permission: ECtHR's Endorsement of Evidence Repurposing*, cit., 2.

Diritto di critica e sanzione detentiva: “*chilling effect*” anche al di fuori della professione giornalistica?*

Alice Tadiotto

Corte di cassazione, sez. V penale, 21 maggio 2025, n. 29840

L'irrogazione di una pena detentiva, ancorché sospesa, per il reato di diffamazione connesso ai mezzi di comunicazione (nella specie, il social network Facebook), anche se non commesso nell'ambito dell'attività prettamente giornalistica, può essere compatibile con la libertà di espressione garantita dall'art. 10 CEDU soltanto in circostanze eccezionali, qualora siano stati lesi gravemente altri diritti fondamentali, come, per esempio, in caso di discorsi di odio o di istigazione alla violenza (nella specie, la Corte ha rilevato che escludere la pena detentiva soltanto nelle ipotesi di diffamazione commessa nell'esercizio dell'attività giornalistica rischierebbe di generare frizioni con il principio di uguaglianza, ex art. 3, c. 1, Cost., e con il principio di ragionevolezza, ex art. 3, c. 2, Cost., prevedendo un trattamento sanzionatorio sfavorevole per fatti di solito connotati da minore gravità e/o diffusività, e dunque complessiva offensività, rispetto a quelli commessi nell'esercizio dell'attività giornalistica).

Sommario

1. Premessa. – 2. Diffamazione e pena detentiva: la più recente evoluzione giurisprudenziale. – 2.1. La giurisprudenza della Corte Edu. – 2.2. L'intervento della Corte costituzionale. – 3. L'iter argomentativo seguito dalla Cassazione. – 4. Pena detentiva e diffamazione non commessa nell'esercizio dell'attività giornalistica... – 5. ... solo in “casi eccezionali”. – 6. Conclusioni.

1. Premessa.

Con questa recente sentenza, la Corte di Cassazione ha preso posizione sul tema della compatibilità del ricorso alla pena detentiva per il reato di diffamazione realizzato mediante mezzi comunicativi di rapida e duratura amplificazione, anche se non commesso nell'ambito di attività giornalistica.

* L'articolo è stato sottoposto, in conformità al regolamento della Rivista, a referaggio anonimo.

La vicenda ha tratto origine dalla pubblicazione, sull'account Facebook dell'imputato, di un post nel quale si accusava la persona offesa – all'epoca dei fatti Presidente di una Corte d'Assise d'Appello siciliana – di trovarsi in situazione di incompatibilità ai sensi dell'art. 19 ordinamento giudiziario in ragione di rapporto di parentela (quale coniuge e madre) di altri magistrati nel medesimo distretto giudiziario e di aver emesso provvedimenti in processi nei quali il coniuge avrebbe svolto funzioni quali rappresentante della pubblica accusa.

Si trattava di affermazioni palesemente non veritiere, atteso che la vittima non era sposata con il magistrato indicato come suo coniuge, né aveva figlie femmine, ma solo maschi, nessuno dei quali apparteneva alla magistratura. Dopo la pubblicazione nell'account Facebook dell'imputato, queste false informazioni sono poi state riprese anche da alcune testate giornalistiche online.

All'esito del processo con conformi giudizi di merito, l'imputato è stato condannato per il reato di diffamazione, alla pena – sospesa – di quattro mesi di reclusione.

Escluso qualunque profilo in ordine alla corretta applicazione della fattispecie di diffamazione, nella specie, la Quinta sezione della Corte di Cassazione ha ritenuto che l'irrogazione della pena detentiva, ancorché sospesa, per il reato di diffamazione connesso ai mezzi di comunicazione, anche se non commesso nell'ambito dell'attività prettamente giornalistica, possa essere compatibile con la libertà di espressione garantita dall'art. 10 CEDU «soltanto in circostanze eccezionali, qualora siano stati lesi gravemente altri diritti fondamentali, come per esempio, in caso di discorsi d'odio o di istigazione alla violenza»¹.

La Corte di Cassazione ha, quindi, ritenuto fondato il motivo di ricorso riferito alla censura di eccessiva gravosità del trattamento sanzionatorio, con speciale riguardo alla necessità di motivare espressamente e puntualmente l'eccezionale gravità della condotta diffamatoria, ed ha rinviato al giudice di merito il compito di accertarne l'eventuale ricorrenza nel caso di specie. La questione della compatibilità tra pena detentiva e reato di diffamazione, intimamente legata al bilanciamento tra onore e libertà di manifestazione del pensiero, è stata oggetto di un nutrito dibattito giurisprudenziale e dottrinale², che è sfociato anche in un noto intervento della Corte costituzionale³.

¹ Cass. pen., sez. V, 21 maggio 2025, n. 29840.

² Sul punto si veda A. Gullo, *Diffamazione e legittimazione dell'intervento penale. Contributo a una riforma dei delitti contro l'onore*, Roma, 2013, 57 ss.

³ V. Corte cost., 9 giugno 2020, n. 132; Corte cost., 12 luglio 2021, n. 150. Per un commento sull'ordinanza si veda M. Cuniberti, *La pena detentiva per la diffamazione tra Corte costituzionale e Corte europea dei diritti dell'uomo: l'ordinanza della Corte costituzionale n. 132 del 2020*, in *Osservatorio costituzionale*, 5, 2020; C. Melzi d'Eril – G. E. Vigevani, *La riforma della diffamazione: da Strasburgo al Senato, passando per Palazzo della Consulta*, in questa *Rivista*, 2020, 3, 137 ss.; C. Magnani, *Diffamazione e pena detentiva: la libertà di informazione tra ordinamento interno e CEDU nella ordinanza 132 del 2020 della Consulta*, in *Forum di Quaderni Costituzionali*, 2, 2021; F. Perchinunno, *Riflessioni a prima lettura sull'ordinanza n. 132/2020 della Corte costituzionale*, in *Federalismi.it*, 27, 2020; A. Ruggeri, *Replicato, seppur in modo più cauto e accorto, alla Consulta lo schema della doppia pronuncia inaugurato in Cappato (nota minima a margine di Corte cost. n. 132 del 2020)*, in *ConsultaOnline*, 2, 2020; R. Pinardi, *La Corte ricorre nuovamente*

L'importanza e la novità della presente decisione si coglie soprattutto nel segnalato riferimento alla necessità di estendere i principi elaborati dalla giurisprudenza convenzionale anche al di fuori di fatti commessi nell'esercizio specifico di attività giornalistica.

2. Diffamazione e pena detentiva: la più recente evoluzione giurisprudenziale

2.1. La giurisprudenza della Corte Edu

Come è noto, l'emergere della questione relativa alla congruità del trattamento sanzionatorio per i fatti di diffamazione e il suo imporsi nella riflessione teorica sono il portato di un percorso che ha visto nel tempo la giurisprudenza nazionale aprirsi ai principi elaborati dalla Corte Europea dei Diritti dell'Uomo⁴.

Le sentenze della Corte di Strasburgo hanno cercato, da un lato, di individuare i confini del legittimo esercizio del diritto di critica politica e giudiziaria⁵; dall'altro hanno avuto ad oggetto, per l'appunto, il trattamento

alla discussa tecnica decisionale inaugurata col caso Cappato, in *Forum di Quaderni Costituzionali*, 3, 2020; D. Casanova, *L'ordinanza n. 132 del 2020 sulla pena detentiva per il reato di diffamazione mezzo stampa: un altro (preoccupante) rinvio della decisione da parte del Giudice costituzionale*, in *ConsultaOnline*, 3, 2020; A. Mazzola, *Decide che deciderà! La Corte costituzionale torna a adoperare la tecnica inaugurata con il "caso Cappato"*, in *ConsultaOnline*, 3, 2020; C. Cipolloni, *La rimodulazione del "punto di equilibrio" tra la libertà di stampa e la tutela della reputazione individuale nella "rete" dell'evoluzione tecnologica: prime osservazioni a margine dell'ordinanza della Corte costituzionale n. 132 del 2020*, in *Dirittifondamentali*, 2, 2021. Per un commento sulla sentenza si veda C. Magnani, *La Corte costituzionale rivede la detenzione per la diffamazione tramite i media. Prime note sulla sentenza n. 150 del 2021*, in *rivista Aic*, 5 ottobre 2021, 5, 2021, 171 ss.; A. Cardone, *Pena detentiva per la diffamazione e funzione democratica della libertà di espressione: quid iuris oltre il caso della professione giornalistica?*, in *Consulta online*, 24 febbraio 2022, 1, 2022, 349 ss.; C. Malavenda, *La sentenza n. 150/2021 della Corte Costituzionale in tema di diffamazione: i "pericoli per la democrazia" e il rischio che l'informazione, da "cane da guardia", si trasformi in "cucciolo da salotto"*, in *Giurisprudenza penale*, 21 luglio 2021; A. Napolitano, *Il difficile bilanciamento tra libertà di informazione professionale e tutela della reputazione della persona. Riflessioni sulla dichiarazione di incostituzionalità della pena detentiva nei confronti dei giornalisti*, in *questa Rivista*, 1, 2022, 272 ss.

⁴ Così A. Gullo, *Diffamazione e pena detentiva, Riflessioni su Cass. Pen., sez. V, 11.12.2013 (13 marzo 2014)*, n. 12203, Pres. Marasca, Rel. Lapalorcia, Ric. Strazzacapa, in *Diritto penale contemporaneo*, 13 marzo 2016.

⁵ Sul tema la Corte di Strasburgo si è occupata, in particolare, di individuare i limiti entro cui è concesso ai giornalisti l'esercizio del diritto di critica, nel bilanciamento tra libertà prevista dall'art. 10 CEDU e tutela dell'onore di cui al par. 2 del medesimo articolo e, della vita privata, di cui all'art. 8 CEDU. Secondo la Corte EDU, la ponderazione di questi due valori si deve attuare differenziando i casi in cui la libertà di stampa abbia ad oggetto questioni politiche e di pubblico interesse, da quelli in cui riguardi la persona di soggetti politici, ai quali si richiede un alto tasso di tolleranza alla critica, soprattutto allorché quest'ultima si inserisca in un contesto di agone politico, ove prevale l'interesse a tenere alto il livello di dibattito pubblico. Il ricorso ad affermazioni esagerate, provocatorie e smodate è ammesso, per la giurisprudenza della Corte EDU, al fine di consentire che il confronto politico, quale "cuore della democrazia" si possa svolgere il più liberamente possibile (così la Corte EDU, nella sentenza: CEDU, *Diugowiecki c. Polonia*, ric. 23806/03 (2009), § 37). In dottrina si veda F. Oriana, *Diffamazione e pena detentiva: in attesa del legislatore*,

sanzionatorio della diffamazione a mezzo stampa, punita con sanzioni severe nella maggior parte degli ordinamenti europei.

A partire dal *leading case* *Cumpănaș e Mazăre c. Romania*⁶, si è affermato nella giurisprudenza della Corte Europea dei Diritti dell'Uomo un indirizzo interpretativo volto a considerare l'irrogazione della pena detentiva un'ingerenza sproporzionata con il diritto alla libertà di espressione dei giornalisti, nel loro ruolo essenziale di "cani da guardia" (*watchdogs*) o "sentinelle" della democrazia. In particolare, secondo la Corte di Strasburgo, la sola previsione astratta della pena detentiva, a prescindere da eventuali sospensioni o commutazioni della sanzione, produce un effetto dissuasivo (c.d. *chilling effect*) che inibisce l'esercizio della libertà di stampa e costituisce una minaccia alla libertà di informazione incompatibile con l'art. 10 CEDU⁸. Le misure statali, infatti, ancorché necessarie per tutelare l'altrui onore e reputazione, potrebbero dissuadere i media dal compito di segnalare all'opinione pubblica casi apparenti o supposti di abuso dei pubblici poteri⁹. Sarebbe, quindi, proprio l'interesse pubblico alla conoscenza dei fatti a dover prevalere sull'esigenza di prevenzione e repressione dei medesimi. In altri termini, secondo i giudici di Strasburgo, anche a fronte dell'espresso riconoscimento della responsabilità penale in capo ai giornalisti, la pena detentiva per i delitti connessi alla sfera della libertà di espressione, così come multe e risarcimenti eccessivi o sanzioni deontologiche sproporzionate, contrastano con l'art. 10 CEDU. Ciò con l'esclusione dei c.d. "casi eccezionali", che la Corte EDU individua nella lesione dei diritti fonda-

dalla cassazione nuovi spunti sul difficile equilibrio fra libertà di manifestazione del pensiero e tutela della reputazione, in *Sistema penale*, 1, 2022, 102.

⁶ CEDU, *Cumpănaș e Mazăre c. Romania*, ric. 33348/96 (2004).

⁷ Risale alla sentenza *Goodwin c. Regno Unito* del 1996 la consapevolezza della qualità di «*public watchdog*» dei mezzi di informazione, dato il loro compito di «segnalare all'opinione pubblica casi apparenti o soltanto supposti di abuso dei pubblici poteri e di svolgere una funzione propulsiva rispetto a questioni di interesse pubblico». V. CEDU, *Goodwin c. Regno Unito*, ric. 17488/90 (1996), § 39; più volte ribadito in CEDU, *Barthold c. Germania*, ric. 8734/79 (1985), § 58; CEDU, *Lingens c. Austria*, n. 9815/82 (1986), § 44; CEDU, *Sunday Times c. Regno Unito* (n. 2), ric. 13166/87 (1991), § 50.

⁸ L'art. 10 al primo paragrafo stabilisce che «ogni persona ha diritto alla libertà d'espressione. Tale diritto include la libertà d'opinione e la libertà di ricevere o di comunicare informazioni o idee senza che vi possa essere ingerenza da parte delle autorità pubbliche e senza limiti di frontiera. Il presente articolo non impedisce agli Stati di sottoporre a un regime di autorizzazione le imprese di radiodiffusione, cinematografiche o televisive». Mentre il secondo comma precisa che: «L'esercizio di queste libertà, poiché comporta doveri e responsabilità, può essere sottoposto alle formalità, condizioni, restrizioni o sanzioni che sono previste dalla legge e che costituiscono misure necessarie, in una società democratica, alla sicurezza nazionale, all'integrità territoriale o alla pubblica sicurezza, alla difesa dell'ordine e alla prevenzione dei reati, alla protezione della salute o della morale, alla protezione della reputazione o dei diritti altrui, per impedire la divulgazione di informazioni riservate o per garantire l'autorità e l'imparzialità del potere giudiziario». La Corte EDU ritiene compatibile con la libertà di espressione sancita dall'art. 10 l'imposizione di sanzioni civili, penali o amministrative per condotte di diffamazione solo se previste dalla legge, giustificate da uno degli scopi legittimi previsti dall'art. 10, par. 2, CEDU e «necessarie in una società democratica» per raggiungere tali scopi (v. CEDU, *Pedersen e Baadsgaard c. Danimarca*, ric. 49017/99 (2004), § 67).

⁹ V. L. Tomasi, *Diffamazione a mezzo stampa e libertà di espressione nell'orizzonte della tutela integrata dei diritti fondamentali*, in *Diritto penale contemporaneo, Rivista Trimestrale*, 4, 2020, 59.

mentali: «l'imposizione di una pena detentiva per un reato di stampa sarà compatibile con la libertà di espressione dei giornalisti come garantita dall'art. 10 della Convenzione solo in circostanze eccezionali, in particolare laddove altri diritti fondamentali siano stati seriamente compromessi, come, ad esempio in caso di discorsi di incitamento all'odio o di istigazione alla violenza»¹⁰. Questi principi si sono consolidati all'interno della giurisprudenza della Corte EDU¹¹ e sono stati ribaditi anche in pronunce che riguardano precipuamente l'ordinamento italiano¹².

Alla base dell'impostazione della giurisprudenza sovranazionale vi è, in sostanza, l'esigenza di assicurare ampi spazi di libertà alla stampa quale veicolo di trasmissione di informazioni ai cittadini. Questi principi sono poi stati ripresi dalla stessa Corte di Strasburgo in casi di critica politica e di critica giudiziaria, rispetto ai quali è stata sottolineata l'esigenza di tenere conto della delicatezza delle funzioni esercitate e del riserbo che dovrebbe connotarle, senza dimenticare che anche in questo fronte la stampa può esercitare un controllo¹³.

Posizioni simili a quelle della Corte EDU sono condivise dal Consiglio europeo, il quale, nell'esprimere la posizione politica dell'Unione Europea, ha ribadito in più occasioni la tendenziale incompatibilità della pena detentiva con la libertà di espressione dei giornalisti, ed ha esortato gli Stati membri ad una totale depenalizzazione della diffamazione¹⁴.

2.2. L'intervento della Corte costituzionale

L'affermazione della tendenziale illegittimità convenzionale della pena de-

¹⁰ CEDU, *Cumpană e Mazăre c. Romania*, cit.

¹¹ *Ex multis*, v. CEDU, *Katrami c. Grecia*, ric. 19331/05 (2007); CEDU, *Kydonis c. Grecia*, ric. 24444/07 (2009); CEDU, *Fatullayev c. Azerbaijan*, ric. 40984/07 (2010); CEDU, *Egeland e Hanseid c. Norvegia*, ric. 34438/04 (2009) e, più di recente, la pronuncia dalla Corte europea dei diritti dell'uomo nel caso *Gheorghe Florin Popescu c. Roumanie*, ric. 79671/13 (2021).

¹² Si veda, in particolare, CEDU, *Belpietro c. Italia*, ric. 42612/10 (2013). Per la ricostruzione della vicenda e un commento sul tema si rimanda a A. Giudici, *Il caso Belpietro c. Italia: la pena detentiva per la diffamazione è contraria all'art. 10 CEDU*, in *Diritto penale contemporaneo*, 26 settembre 2013 e C. Melzi d'Eril, *La Corte Europea condanna l'Italia per sanzione e risarcimento eccessivi in un caso di diffamazione. Dalla sentenza qualche indicazione per la magistratura, il legislatore e le parti?*, in *Diritto penale contemporaneo*, 12 novembre 2012; CEDU, *Ricci c. Italia*, ric. n. 30210/06 (2013), per la ricostruzione della vicenda si rimanda a S. Rossetti, *La corte EDU sul bilanciamento tra riservatezza delle comunicazioni e libertà di espressione del giornalista*, in *Diritto penale contemporaneo*, 5 novembre 2013. CEDU, *Sallusti c. Italia*, ric. 22350/13 (2019), per un commento si rimanda a S. Turchetti, *Diffamazione, pena detentiva, caso Sallusti: ancora una condanna*, in *Diritto penale contemporaneo*, 18 marzo 2019; S. Lonati, *Diffamazione a mezzo stampa e applicazione della pena detentiva: ancora qualche riflessione a margine del cd. caso Sallusti in (perenne) attesa di un intervento del legislatore*, in questa Rivista, 1, 2020, 69 ss.

¹³ Cfr. CEDU, *Morice c. Francia*, ric. 29369/10 (2015); CEDU, *L.P. Carvalho c. Portogallo*, ric. 24845/13 49103/15 (2019).

¹⁴ *Resolution 1577/2007 "Towards decriminalisation of defamation"*; *Recommendation 1589/2003 "On freedom of expression in the media in Europe"*; *Resolution 1920/2013 "The state of freedom in Europe"*; *Resolution 2035/2015 "Protection of the safety of journalists and of media freedom in Europe"*; *Resolution 2141/2017 "Attacks against journalists and media freedom in Europe"*.

tentiva ha avuto ripercussioni anche all'interno dell'ordinamento nazionale portando, in assenza di un intervento legislativo in tema¹⁵, dapprima ad un adeguamento giurisprudenziale¹⁶ e poi all'intervento della Corte costituzionale¹⁷.

Come già segnalato, la Consulta si è pronunciata su questi temi, prima con l'ordinanza n. 132 del 2020 e, successivamente, con la sentenza n. 150 del 2021, recependo, nella sostanza, le indicazioni provenienti dalla Corte EDU.

Nella specie, il quesito sulla compatibilità costituzionale e convenzionale della previsione di pene detentive per la diffamazione a mezzo stampa è stato sollevato con riferimento all'art. 13, legge 8 febbraio 1948, n. 47 (recante le «Disposizioni sulla stampa»)¹⁸, ed in relazione all'art. 595, c. 3, c.p. Pur riconoscendo la necessità e urgenza di una complessiva rimeditazione del bilanciamento tra libertà di manifestazione del pensiero del giornalista e tutela della reputazione individuale, la Corte costituzionale, con l'ordi-

¹⁵ Per una completa ricostruzione del trattamento sanzionatorio della diffamazione a mezzo stampa e le varie proposte di riforma si veda M. Pisapia – C. Cherchi, *Detenzione e libertà di espressione. Riflessioni sul trattamento sanzionatorio del reato di diffamazione a mezzo stampa in occasione della pronuncia della Corte Costituzionale*, in *Giurisprudenza Penale Trimestrale*, 2020, 100 ss.

¹⁶ In questo senso si veda, ad esempio, Cass. pen., sez. V, 11 dicembre 2013, n. 12203, relativa alla responsabilità del direttore per omesso controllo sul contenuto di una pubblicazione diffamatoria (art. 57 cod. pen.), secondo cui non si giustifica, per la CEDU, l'applicazione della pena detentiva, nemmeno in caso di pubblicazione di una notizia falsa; in termini simili v. anche Cass. pen., sez. V, 19 novembre 2015, n. 4298; Cass. pen., Sez. V, 10 luglio 2019, n. 38721; Cass. pen., sez. V, 9 luglio 2020, n. 26509.

¹⁷ La dottrina da tempo aveva segnalato l'opportunità di un intervento da parte della Corte costituzionale. In particolare si veda A. Giudici, *Il caso Belpietro c. Italia: la pena detentiva per la diffamazione è contraria all'art. 10 CEDU*, cit., 2013; C. Melzi d'Eril, *La Corte europea condanna l'Italia per sanzione e risarcimento eccessivi in un caso di diffamazione*, cit.; M. Cuniberti, *Pene detentive per la diffamazione, responsabilità del direttore e insindacabilità delle opinioni del parlamentare: il "caso Belpietro" davanti alla Corte europea dei diritti dell'uomo*, in *Osservatorio costituzionale*, 2014; M. Bernardini, *Ancora sul trattamento sanzionatorio del reato di diffamazione. Il caso Sallusti al vaglio della Corte Edu*, in *Diritti comparati*, 19 giugno 2019; S. Lonati, *Diffamazione a mezzo stampa e applicazione della pena detentiva*, cit; A. Tarallo, *L'esimente del diritto di cronaca nella prospettiva della Cedu: nota alla sentenza Magosso-Brindani contro Italia*, in *Dirittifondammetali.it*, 1, 2020.

¹⁸ L'art. 13 della l. n. 47 del 1948 sanziona(va) con la reclusione da uno a sei anni e con la multa non inferiore a euro 258 la diffamazione commessa con il mezzo della stampa consistente nell'attribuzione di un fatto determinato. La questione di costituzionalità riguardava, in particolare, il contrasto tra il diritto interno e il diritto della CEDU assunto a parametro interposto di costituzionalità tramite l'art. 117, primo comma, cost. Si vedano al riguardo le c.d. "sentenze gemelle" della Corte costituzionale nn. 348 e 349 del 2007; in particolare con l'ordinanza iscritta al n. 140 del r.o. 2019 il Tribunale ordinario di Salerno, sez. II penale, ha sollevato, in riferimento agli artt. 3, 21, 25, 27 e 117, c. 1, Cost., quest'ultimo in relazione all'art. 10 CEDU, questioni di legittimità costituzionale dell'art. 595, c. 3, c.p. e dell'art. 13 della legge 8 febbraio 1948, n. 47 (Disposizioni sulla stampa). Con l'ordinanza iscritta al n. 149 del r.o. 2019 il Tribunale ordinario di Bari, sezione prima penale, ha sollevato, in riferimento all'art. 117, c. 1, Cost., in relazione all'art. 10 CEDU, questione di legittimità costituzionale dell'art. 13 della legge n. 47 del 1948, «in combinato disposto» con l'art. 595 cod. pen., «nella parte in cui sanziona il delitto di diffamazione aggravata, commessa a mezzo stampa e consistente nell'attribuzione di un fatto determinato, con la pena cumulativa della reclusione da uno a sei anni e della multa non inferiore a 256 [recte: 258] euro, invece che in via alternativa».

nanza n. 132 del 2020 ha rinviato la trattazione delle questioni all'udienza pubblica del 22 giugno 2021, per consentire al legislatore di intervenire con una nuova disciplina della materia.

In assenza di intervento legislativo, preso atto che il bilanciamento tra tutela della reputazione e libertà di informazione è divenuto ormai inadeguato, con la sentenza n. 150 del 2021 la Corte costituzionale ha dichiarato l'incostituzionalità dell'art. 13 della legge n. 47 del 1948, mentre ha rigettato la questione di legittimità riguardante l'art. 595 c.p., ritenendo possibile un'interpretazione costituzionalmente e convenzionalmente orientata della norma scrutinata.

Con riferimento al primo quesito sollevato, la Corte costituzionale ha concluso che l'indefettibile applicazione della pena detentiva prevista dall'art. 13 della legge n. 47 del 1948 è divenuta ormai incompatibile con l'esigenza «di non dissuadere per effetto del timore della sanzione privativa della libertà personale, la generalità dei giornalisti dall'esercitare la propria cruciale funzione di controllo sull'operato dei pubblici poteri»¹⁹.

La Consulta ha, invece, rigettato la questione di legittimità riguardante l'art. 595 c.p., rilevando come la (eventuale) applicazione della pena detentiva in caso di diffamazione compiuta a mezzo della stampa o di altro mezzo di pubblicità non sia di per sé incompatibile con le ragioni di tutela della libertà di manifestazione del pensiero, specie nel caso in cui la diffamazione si caratterizzi in termini di eccezionale gravità.

Nonostante l'esplicito riconoscimento della libertà di espressione quale «pietra angolare dell'ordine democratico»²⁰, la Corte costituzionale ha ritenuto la reputazione individuale un diritto del pari inviolabile, strettamente legato alla stessa dignità della persona di cui agli artt. 2 e 3 Cost. e tutelato dalla stessa Corte Europea dei Diritti dell'Uomo²¹. Tale diritto può essere oggetto di aggressioni illegittime fortemente lesive della dignità personale, idonee ad incidere pesantemente sulla vita privata, familiare e professionale. In tempi recenti, inoltre, il carattere lesivo per la vittima di diffamazione risulta ulteriormente potenziato dalla rapida e duratura amplificazione degli addebiti diffamatori determinata dai social networks e dai motori di ricerca in internet.

Questi effetti negativi devono essere prevenuti dall'ordinamento con strumenti idonei e proporzionati, tra i quali, secondo la Consulta, non può in assoluto escludersi il ricorso alla sanzione detentiva, la cui applicazione deve, però, essere circoscritta a situazioni di «eccezionale gravità»²² e cir-

¹⁹ Corte cost., 12 luglio 2021, n. 150, punto 4.2.

²⁰ La Corte costituzionale si richiama a una propria precedente sentenza in tema la n. 84 del 1969. A ulteriore testimonianza della centralità della libertà di manifestazione del pensiero nell'ordinamento costituzionale l'ordinanza n. 132 del 2020 ha ricordato che la prima sentenza della Corte costituzionale (Corte cost., 14 giugno 1956, n. 1) ha dichiarato l'incostituzionalità di una norma (art. 2 T.U.L.P.S.) proprio per contrasto con l'art. 21 Cost.

²¹ In particolare, la CEDU tutela il diritto alla reputazione individuale attraverso la garanzia del «Diritto al rispetto della vita privata e familiare» di cui all'art. 8 CEDU.

²² Corte cost., 12 luglio 2021, n. 150, punto 6.2 del considerato in diritto. La Corte costituzionale si richiama espressamente alle altre sentenze della Corte EDU in tema: CEDU, *Cumpănă e Mazăre c. Romania*, cit., § 115; nonché sentenze, CEDU, *Balaskas c. Grecia*, ric. 73087/17 (2020), § 61; *Atamanchuk c. Russia*, ric. 4493/11 (2020), § 67; CEDU, *Sallusti c. Italia*, cit., § 59; CEDU, *Belpietro c. Italia*, cit., § 53; CEDU, *Katrami c. Grecia*, cit., § 39.

condana «da cautele idonee a schermare il rischio di indebita intimidazione esercitato su chi svolge la professione giornalistica»²³.

Per individuare tali situazioni di particolare gravità la Corte costituzionale si richiama, ancora una volta, all'elaborazione della giurisprudenza di Strasburgo. Secondo la CEDU, come visto, simili ipotesi eccezionali risultano integrate in presenza di discorsi d'odio e di istigazione alla violenza. La Corte costituzionale aggiunge che casi egualmente eccezionali, tali da giustificare l'inflizione di sanzioni detentive, potrebbero, ad esempio, essere rappresentati anche da campagne di disinformazione condotte attraverso la stampa, internet o i social media, caratterizzate dalla diffusione di addebiti gravemente lesivi della reputazione della vittima, e compiute nella consapevolezza da parte dei loro autori della – oggettiva e dimostrabile – falsità degli addebiti stessi.

Chi pone in essere simili condotte – eserciti o meno la professione giornalistica – «certo non svolge la funzione di “cane da guardia” della democrazia», ma, all'opposto, «crea un pericolo per la democrazia, combattendo l'avversario mediante la menzogna, utilizzata come strumento per screditare la sua persona agli occhi della pubblica opinione»²⁴.

Al di fuori delle situazioni di eccezionale gravità (che giustificano il ricorso alla sanzione detentiva), la prospettiva del carcere deve rimanere esclusa, secondo la Consulta, per il giornalista, così come per chiunque altro abbia manifestato attraverso la stampa o altri mezzi di pubblicità la propria opinione, restando aperta solamente la possibilità di ricorrere a rimedi e/o sanzioni diverse da quella detentiva.

In definitiva, quindi, il carattere alternativo della sanzione detentiva previsto dall'art. 595, c. 3, c.p. viene considerato dalla Corte costituzionale compatibile con l'art. 21 Cost e con l'art. 10 CEDU²⁵.

3. L'iter argomentativo seguito dalla Cassazione

In questo panorama giurisprudenziale si inserisce la recente pronuncia della Corte di Cassazione la quale, dando conto dell'evoluzione interpretativa sopra richiamata, estende i principi elaborati dalla Corte EDU e dalla Corte costituzionale anche alle condotte di diffamazione connesse ai mezzi di comunicazione (nella specie il social network Facebook), anche se non commessi nell'esercizio dell'attività giornalistica.

Nella specie, la Corte ritiene che «l'irrogazione di una pena detentiva, ancorché sospesa, per il reato di diffamazione connesso ai mezzi di comu-

²³ Corte cost., 12 luglio 2021, n. 150, punto 6.2.

²⁴ *Ibid.* Secondo la Corte, inoltre, se circoscritta ai casi sopra richiamati, la previsione astratta e la concreta applicazione di sanzioni detentive non possono produrre effetti di indebita intimidazione nei confronti dell'esercizio della professione giornalistica.

²⁵ La sentenza n. 150 del 2020 contiene anche un ulteriore dispositivo che esula dalle disposizioni direttamente oggetto delle questioni di costituzionalità. La Corte ha dichiarato l'illegittimità costituzionale dell'art. 30, quarto comma, della legge n. 223 del 1990. Tale comma venne introdotto per rimediare alla disparità di trattamento fra stampa e televisione, ma come rilevato in dottrina (C. Malavenda – C. Melzi d'Eril – G. E. Vigevani, *Le regole dei giornalisti*, Bologna, 2021, 75), ha prodotto un nuovo squilibrio.

nicazione, anche se non commesso nell'ambito dell'attività prettamente giornalistica, possa essere compatibile con la libertà di espressione garantita dall'art. 10 CEDU soltanto in circostanze eccezionali, qualora siano stati lesi gravemente altri diritto fondamentali, come, per esempio, in caso di discorsi di odio o di istigazione alla violenza»²⁶.

Per giungere a tale conclusione la Corte di Cassazione si è, innanzitutto, richiamata alla giurisprudenza della Corte EDU secondo la quale, come visto, l'irrogazione della pena detentiva, ancorché sospesa, per il delitto di diffamazione commesso mediante mezzi comunicativi di rapida e duratura amplificazione, deve essere connessa alla grave lesione di diritti fondamentali, come nel caso di discorsi di odio o di istigazione alla violenza.

La Corte di Cassazione riconosce che la giurisprudenza della Corte EDU, è stata elaborata con precipuo riferimento alla proporzione della pena detentiva «nell'ambito dell'esercizio del diritto di cronaca e di critica giornalistica, per l'effetto dissuasivo (c.d. chilling effect) che può determinare sulla libertà di espressione della stampa in generale, considerato il “*match-dog*” della democrazia»²⁷. Nota, tuttavia, che la stessa giurisprudenza di Strasburgo ha attribuito sistematicamente rilievo «anche al rischio di effetto dissuasivo rispetto alla libertà di manifestazione del pensiero critico anche in relazione all'esercizio del diritto di critica non connesso con la libertà di stampa, che sotto il profilo della natura e della severità della sanzione che fa da contrappeso alla condotta di critica diffamatoria, si traduce in un monito a limitare la necessità della pena detentiva ai soli casi eccezionali»²⁸. La Corte di Cassazione cita in particolare i precedenti della Corte EDU, *Gran Camera, Morice c. Francia*, del 23 aprile 2015 e Corte EDU, *L.P. Carvalho c. Portogallo*, del 8 ottobre 2019 nei quali, ricorda, «la vicenda riguardava proprio un caso di critica diretta nei confronti degli organi giudiziari»²⁹. Per altro verso, la Corte di legittimità si richiama alle recenti pronunce della Corte costituzionale (ord. n. 131 del 2020 e sent. n. 150 del 2021) in forza delle quali, ove non sia motivata la situazione eccezionale connessa alla grave lesione dei diritti fondamentali che l'ordinamento pone in bilanciamento con il diritto alla libera manifestazione del pensiero, la determinazione della pena detentiva non è costituzionalmente giustificata.

A questo punto la Corte di Cassazione fa riferimento alla propria giurisprudenza che in una recente sentenza (Cass. pen., sez. V, 17 febbraio 2021, n. 13993, Scaffidi, Rv. 281024 – 01) ha esteso i principi sopra richiamati dalla diffamazione a mezzo stampa a qualsiasi altra forma di diffamazione, ritenendo che gli stessi valgano anche qualora si sia verificata una rapida e duratura amplificazione degli addebiti diffamatori determinata dai social networks.

Da ultimo, escludere la pena detentiva – riservandola soltanto ai c.d. discorsi d'odio – soltanto nelle ipotesi di diffamazione commessa nell'esercizio dell'attività giornalistica, secondo la Cassazione rischierebbe di generare frizioni con il principio di uguaglianza (art. 3, c. 1, Cost.) e con il

²⁶ Cass. pen., sez. V, 21 maggio 2025, n. 29840.

²⁷ *Ibid.*

²⁸ *Ibid.*

²⁹ *Ibid.*

principio di ragionevolezza (art. 3, c. 2, Cost.), prevedendo un trattamento sanzionatorio particolarmente affittivo (la pena detentiva) per fatti di solito connotati da minore gravità e/o diffusività, e dunque complessiva offensività, rispetto a quelli commessi nell'esercizio dell'attività giornalistica.

4. Pena detentiva e diffamazione non commessa nell'esercizio dell'attività giornalistica...

La sentenza in esame si inserisce in quel recentissimo filone della giurisprudenza di legittimità che estende i principi elaborati dalla Corte di Strasburgo (e richiamati quasi integralmente nelle pronunce della Corte costituzionale) anche alle condotte diffamatorie connesse ai mezzi di comunicazione, ancorché non commesse nell'ambito dell'attività prettamente giornalistica. -

Per addivenire a questa conclusione, come anticipato, la Cassazione si è richiamata ad alcune sentenze della Corte EDU che, a prima vista, potrebbero sollevare qualche perplessità in quanto riferite a situazioni parzialmente diverse rispetto a quella che ha portato alla pronuncia in commento. In particolare, sia nel caso *Morice c. Francia*³⁰ che *L.P. Carvalho c. Portogallo*³¹,

³⁰ CEDU, *Morice c. Francia*, cit. In breve, la vicenda: la ricorrente era un avvocato che, nel corso degli anni novanta del secolo scorso, ha assistito la moglie di un giudice francese, Borrel, ufficialmente morto per suicidio in un Paese africano, ma di cui in realtà si sospetta l'omicidio. Disposta l'apertura di una nuova inchiesta, è stata incaricata quale giudice istruttore M. I ritardi nella conduzione dell'inchiesta hanno spinto la Corte d'Appello ad attribuire l'affare a un altro giudice istruttore. In questo contesto, il 4 luglio del 2000, - J.M. - esprime pubblicamente solidarietà e fiducia alla collega M., al termine di un'assemblea generale dei magistrati appartenenti al Tribunale di grande istanza di Parigi; le dichiarazioni d'apprezzamento sono nate dalle notizie di stampa, dove si è riportato che il Consiglio superiore della magistratura avesse aperto - su iniziativa del Ministro di giustizia - un dossier relativo a M. in merito alla correttezza del suo operato nell'ambito di un'indagine su Scientology. A sua volta, il 6 settembre del 2000, la ricorrente ha chiesto al Ministro di giustizia l'avvio di un procedimento disciplinare nei confronti di M. per la parzialità dimostrata nella conduzione dell'inchiesta sulle cause della morte di Borrel. Il giorno dopo, su *Le Monde*, appare un articolo dedicato a tale vicenda, in cui la ricorrente denuncia l'assenza d'imparzialità del giudice M. e la sua connivenza con le autorità giudiziarie del Paese africano. Di qui, il procedimento penale per diffamazione a mezzo stampa contro la ricorrente, il direttore del giornale e il giornalista, giudizio che si conclude con la loro condanna. In appello, la sentenza viene confermata. La ricorrente si rivolge a Strasburgo lamentando la violazione degli art. 6, c. 1, e 10 CEDU. Quanto alla supposta ingerenza nella libertà di espressione, la Corte europea esclude la violazione del dettato convenzionale: l'articolo relativo al giudice M., in breve, è caratterizzato da una particolare virulenza; la ricorrente, inoltre, è stata condannata a una pena pecuniaria, proporzionata rispetto ai fatti in causa. Ha così riassunto la vicenda F. Zacchè, *Monitoraggio Corte EDU luglio-agosto 2013*, in *Diritto penale contemporaneo*, 11 ottobre 2013.

³¹ CEDU, *L.P. e Carvalho c. Portogallo*, cit; In questa sentenza la Corte EDU ha ravvisato una violazione dell'art. 10 della CEDU nel caso di condanna al risarcimento del danno per diffamazione disposta nei confronti di un avvocato per aver criticato, nell'ambito del mandato difensivo attribuitogli, l'operato di un giudice. Alla luce del particolare status soggettivo attribuito agli avvocati, le sentenze di condanna dei giudici nazionali rappresentano un'ingerenza sproporzionata e non necessaria nell'esercizio della libertà di espressione dei professionisti e non corrispondono ad alcun bisogno sociale imperativo. Inoltre, le sanzioni illegittimamente applicate dalle autorità nazionali rischiano di produrre

la Corte EDU si è espressa con riferimento ai limiti del diritto di critica degli avvocati nei confronti dell'operato di un magistrato. In entrambe le sentenze, la Corte di Strasburgo ha posto l'accento sul ruolo centrale che l'avvocato riveste sia nell'amministrazione della giustizia sia nella funzione di intermediazione tra i consociati e gli organi giurisdizionali, contribuendo anche al mantenimento della fiducia dei primi verso i secondi. Nell'esercizio di tale ruolo viene riconosciuto all'avvocato il diritto di critica rispetto all'operato della magistratura e più in generale al funzionamento della giustizia; diritto che può essere esercitato con ampio margine sia all'interno che all'esterno dell'udienza, fermi restando i limiti di carattere deontologico fissati da CCBE (Council of Bars and Law Societies of Europe)³².

Come rilevato da attenta dottrina³³, in entrambi i casi la Corte EDU ha fatto riferimento a soggetti appartenenti a una particolare categoria professionale – avvocati – della quale è apparso necessario salvaguardare l'indipendenza. Quest'ultima potrebbe essere, infatti, compromessa dall'effetto dissuasivo derivante dall'inflizione della pena detentiva per condotte diffamatorie poste in essere nell'esercizio della professione o comunque nell'ambito di un dibattito di interesse generale sul funzionamento della giustizia. Analoghe esigenze non sembrerebbero esserci, almeno *prima facie*, nel caso di specie in cui a venire in rilievo è la critica svolta da un privato cittadino nei confronti di un membro della magistratura.

Se è vero che le sentenze della CEDU sono sempre legate al caso specifico e che è difficile trarre da esse norme generali, non si può trascurare che in quasi tutte le occasioni in cui la Corte EDU si è occupata di pene detentive irrogate in casi di diffamazione o comunque di reati a mezzo stampa ha rinvenuto la sussistenza di una violazione dell'art. 10 CEDU³⁴.

Di ciò appare consapevole anche la Corte costituzionale che, richiamandosi ai principi espressi dalla giurisprudenza convenzionale in materia, ha rilevato come il bilanciamento operato dalla disciplina nazionale tra tutela della reputazione e libertà di informazione appaia ormai inadeguato³⁵.

In conformità alle indicazioni provenienti dal diritto sovranazionale, la Corte costituzionale ha ridimensionato il ricorso alla pena detentiva come mezzo di risposta alla diffamazione sui media. Secondo la Consulta non vi è, però, assoluta incompatibilità costituzionale tra il carcere e alcuni tipi di condotte diffamatorie. Invero, talune forme gravi di diffamazione possono trovare ancora come risposta punitiva la detenzione³⁶.

Questi principi, afferma la Corte costituzionale, devono valere non solo

un effetto deterrente sull'intera professione legale, considerando, in particolare, che le dichiarazioni contestate agli avvocati si inserivano nell'alveo della difesa degli interessi degli assistiti. Per un commento si veda F. Ceci, *Il bilanciamento tra libertà di espressione degli avvocati e tutela dell'autorità del potere giudiziario: un inquadramento sistematico degli interessi nella prospettiva della Corte europea dei diritti dell'uomo*, in questa *Rivista*, 2, 2020, 323 ss.

³² CEDU, *Morice c. Francia*, cit., §§ 132-139.

³³ L. Tomasi, *Diffamazione e illegittimità "convenzionale" della pena detentiva*, cit.

³⁴ In questi termini C. Melzi d'Eril, *La Corte europea condanna l'Italia per sanzione e risarcimento eccessivi in un caso di diffamazione*, cit., 8 ss.

³⁵ Corte cost., 9 giugno 2020, n. 132, punto 6.

³⁶ Corte cost., 12 luglio 2021, n. 150, punto 6.2.

per i professionisti dell'informazione, ma per tutti i cittadini, atteso il ruolo cruciale che la libertà di manifestazione del pensiero riveste nell'ordinamento costituzionale, quale «pietra angolare di ogni ordinamento democratico»³⁷.

In altri termini, secondo la Corte costituzionale, tanto l'attività giornalistica quanto l'espressione del pensiero del *quisque de populo* sono dotate della medesima rilevanza giuridica e appaiono meritevoli del medesimo trattamento sanzionatorio³⁸.

Con riferimento a questo passaggio non sono mancate le critiche di una parte della dottrina³⁹ che ritiene non del tutto ragionevole l'equiparazione, ai fini del regime sanzionatorio, tra la posizione del giornalista professionista e quella del *quisque de populo* che leda l'onore e la reputazione altrui effettuando commenti sui social network o scrivendo su blog. Potrebbe apparire opportuno, secondo questa dottrina, differenziare la posizione del giornalista professionista che con la sua attività contribuisce ad alimentare il pluralismo delle fonti di informazione, dalle dichiarazioni del *quisque de populo* che diffama in «maniera incontenente e randomizzata»⁴⁰.

³⁷ *Ibid.* Nell'ordinanza n. 132 del 2020, la Corte costituzionale ricorda altresì la significativa circostanza storica che la Corte costituzionale – in risposta a ben trenta ordinanze sollevate da giudici comuni – abbia dichiarato l'illegittimità costituzionale di una disposizione di legge proprio in ragione del suo contrasto con l'art. 21 Cost. (sentenza n. 1 del 1956). Sempre nella medesima ordinanza la Corte si riferisce alla libertà di stampa come «coessenziale al regime di libertà garantito dalla Costituzione» (sentenza n. 11 del 1968), «pietra angolare dell'ordine democratico» (sentenza n. 84 del 1969), «cardine di democrazia nell'ordinamento generale» (sentenza n. 126 del 1985 e, di recente, sentenza n. 206 del 2019).

³⁸ Così C. Magnani, *La Corte costituzionale rivede la detenzione per la diffamazione tramite i media*, cit., 191. Nell'espressione di giudizi di valore, invece, come rilevato in dottrina sembra che la Corte EDU proceda a imporre limiti maggiori per il privato cittadino rispetto a una più ampia libertà concessa al giornalista. Per i giornalisti, infatti, con particolare riguardo ai giudizi di valore, la Corte non fa riferimento alla necessità di una base fattuale con riguardo ai giudizi espressi, anche attraverso un linguaggio provocatorio, mentre nel caso del medico che aveva criticato un politico, pur ribadendo l'ampia libertà in un giudizio di valore che non si presta, per sua natura, all'esistenza di una prova, la Corte ha richiesto l'esistenza di una base fattuale che mostri una «sufficiently close connection» con l'opinione espressa. Così M. Castellaneta, *Tutela della reputazione dei politici, libertà di stampa e giudizi di valore con linguaggio provocatorio: la Corte europea procede a un bilanciamento a favore dei giornalisti*, in questa *Rivista*, 1, 2020, 237 ss.

³⁹ A. Cardone, *Pena detentiva per la diffamazione e funzione democratica della libertà di espressione*, cit., 354.

⁴⁰ In queste circostanze non pare che la libertà di espressione possa a tal punto prevalere sulla stessa dignità della persona, tutelata sia dagli artt. 2 e 3 Cost. che dall'art. 8 CEDU, da far ritenere che la pena detentiva possa essere applicata solo nelle circostanze eccezionali rappresentate dall'*hate speech*, dall'istigazione alla violenza e dalle campagne di disinformazione. Per il *quisque de populo* potrebbe essere appropriato continuare a sentire la pressione dell'ordinamento sulle sue modalità espressive e sui contenuti delle sue affermazioni, «non perché questi debbano essere inibiti o repressi, ma perché non c'è motivo che la tutela dell'onore e della reputazione di chi può subire le aggressioni arretri in mancanza della giustificazione democratica che sorregge il diverso bilanciamento che deve essere riservato alla professione giornalistica e, più in generale, a chi opera nelle fonti di informazione». Nei due contesti, quindi, sembrerebbero venire in rilievo diverse esigenze general-preventive e special-preventive e diverso dovrebbe essere il regime sanzionarlo applicato. In questo senso A. Cardone, *Pena detentiva per la diffamazione e funzione democratica della libertà di espressione*, cit., 356-358.

A ben vedere, però, è la stessa Corte costituzionale a riconoscere sia al *quisque de populo* sia al giornalista professionista un ruolo centrale all'interno del dibattito democratico e a precisare che, al di fuori di "casi eccezionali", «la prospettiva del carcere resterà esclusa per il giornalista, così come per chiunque altro che abbia manifestato attraverso la stampa o altri mezzi di pubblicità la propria opinione»⁴¹.

Dello stesso avviso è la Corte di Cassazione nella sentenza in commento. Escludere la pena detentiva, riservandola ai casi eccezionali, per le sole ipotesi di diffamazione commessa nell'esercizio dell'attività giornalistica, rischierebbe, infatti, di generare frizioni sia con il principio di uguaglianza (art. 3, c. 1, Cost.) sia con quello di ragionevolezza (art. 3, c. 2, Cost.).

Da questo punto di vista, per altro verso, potrebbe apparire sicuramente dotata di maggior disvalore e offensività proprio la condotta del giornalista, che diffonda per il tramite della stampa notizie diffamatorie, rispetto a quella del comune cittadino. Il giornalista, infatti, per la natura dell'attività esercitata e i connessi doveri di deontologia professionale, è un soggetto qualificato, da cui deve pretendersi (e attendersi) un maggior grado di diligenza nell'esercitare la propria libertà di espressione senza ledere l'altrui onore e la reputazione. Non pare, allora, potersi giustificare un trattamento sanzionatorio meno sfavorevole nei confronti del giornalista che faccia del suo mestiere "un'arma impropria" per delegittimare altri soggetti.

È vero che l'argomento può essere rovesciato⁴² e il differente trattamento sanzionatorio potrebbe apparire giustificato proprio alla luce del medesimo art. 3 Cost. in forza del peculiare ruolo assunto dalla stampa nella società democratica. Al diritto del giornalista di informare, corrisponde un correlato «diritto all'informazione dei cittadini» e non vi è dubbio che l'attività del giornalista meriti di essere «salvaguardata contro ogni minaccia o coartazione, diretta o indiretta»⁴³ che possa indebolire la sua vitale funzione nel sistema democratico, «ponendo indebiti ostacoli al legittimo svolgimento del suo ruolo di informare i consociati e di contribuire alla formazione degli orientamenti della pubblica opinione, anche attraverso la critica aspra e polemica delle condotte di chi detenga posizioni di potere»⁴⁴. In sostanza, si potrebbe sostenere che è proprio l'esercizio della professione giornalistica a rappresentare l'archetipo della vigilanza democratica che la tutela costituzionale della libertà di manifestazione del pensiero consente di esercitare sui detentori del pubblico potere⁴⁵.

Queste obiezioni non sono, tuttavia, dirimenti e non consentono di superare il contrasto che si verrebbe a creare con l'art. 3 Cost., sia sotto il profilo dell'uguaglianza che della ragionevolezza, se condotte connotate da maggiore offensività venissero sanzionate meno severamente.

Del resto, una notizia falsa, diffusa del giornalista, nel suo ruolo di "cane

⁴¹ V. Corte cost., 12 luglio 2021, n. 150, punto 6.2.

⁴² Come puntualmente rilevato da L. Tomasi, *Diffamazione e illegittimità "convenzionale" della pena detentiva: oltre l'aggravante dell'uso della stampa?*, cit.

⁴³ V. Corte cost., 12 dicembre 1972, n. 172.

⁴⁴ V. Corte cost., 9 giugno 2020, n. 132, punto 7.1.

⁴⁵ Così A. Cardone, *Pena detentiva per la diffamazione e funzione democratica della libertà di espressione*, cit., 355.

da guardia” della democrazia, ha una portata offensiva maggiore anche per l’affidamento che i destinatari ripongono sulla veridicità della stessa e certo non contribuisce ad alimentare il pluralismo delle fonti di informazione ed il dibattito democratico, tramite la ricerca e la pubblicazione di verità “scomode”.

Sarebbe del tutto irragionevole riservare a condotte diffamatorie – anche se consistenti in critiche aspre e taglienti – poste in essere tramite i social network un trattamento maggiormente sfavorevole solo perché non commesse nell’esercizio dell’attività giornalistica.

Se, al di fuori di casi eccezionali, deve escludersi il ricorso alla pena detentiva per il giornalista che travalichi i limiti del legittimo esercizio del diritto di critica, analoga conclusione si impone anche nel caso del comune cittadino. Ed infatti, tanto il giornalista quanto il cittadino nell’esercizio della libertà di espressione sono tenuti al rispetto dei limiti della veridicità dei fatti, della pertinenza degli argomenti e della continenza espressiva⁴⁶.

Di contro, nel caso in cui vengano travalicati i limiti del legittimo esercizio del diritto di critica e di cronaca, tanto per il cittadino quanto per il giornalista, resta «aperta soltanto la possibilità che siano applicate pene diverse dalla reclusione, nonché rimedi e sanzioni civili o disciplinari»⁴⁷. Sempre, come detto, che non si tratti di casi eccezionali.

5. ... solo in “casi eccezionali”

Rimane a questo punto da comprendere quali possano essere i “casi eccezionali” che giustificano il ricorso alla pena detentiva.

Come visto, la Corte EDU nell’enucleare le “ipotesi eccezionali” ha sempre utilizzato una formula generica che richiama i discorsi d’odio e l’istigazione alla violenza lesivi dei diritti fondamentali. Non è mai stato chiarito, tuttavia, se tale elencazione individui in maniera tassativa le ipotesi che giustificano il ricorso alla pena detentiva oppure valga esclusivamente a titolo esemplificativo⁴⁸.

In quest’ultimo senso è orientata da tempo la giurisprudenza di legittimità, all’interno della quale si è consolidato un indirizzo interpretativo volto ad includere casi particolarmente gravi di diffamazione tra quelle ipotesi eccezionali che possono, in concreto, giustificare l’irrogazione della pena detentiva⁴⁹. Più nel dettaglio, la Corte di Cassazione ha, ad esempio, considerato legittimo il ricorso alla pena detentiva in ipotesi di «calcolata e calibrata alterazione della verità dei fatti, finalizzata a disinformare, a creare inesistenti responsabilità e a infliggere fantasiose condanne agli avversari»⁵⁰. Altra “ipotesi eccezionale” è stata considerata la pubblicazione di una

⁴⁶ V. *ex multis*, Cass. pen., sez. V, 5 aprile 2017, n. 31079.

⁴⁷ Così Corte cost., 12 luglio 2021, n. 150, punto 6.2.

⁴⁸ In questo senso M. Cuniberti, *Pene detentive per la diffamazione, responsabilità del direttore e insindacabilità delle opinioni del parlamentare*, cit., 125.

⁴⁹ V. *ex multis*, Cass. pen., sez. V, 26 settembre 2012, n. 41249; Cass. pen., sez. V, 13 marzo 2014, n. 12203; Cass. pen., sez. V, 19 settembre 2019, n. 38721.

⁵⁰ Cass. pen., sez. V, 26 settembre 2012, n. 41249.

notizia non rispondente al vero, nella piena consapevolezza di tale falsità da parte del giornalista⁵¹.

Più incerta è, invece, la posizione della dottrina che, in alcune occasioni, non ha mancato di segnalare l'anomalia della (astratta) previsione della detenzione per il reato di diffamazione a mezzo stampa, ritenendo che "circostanze eccezionali" debbano necessariamente corrispondere a diverse e, più gravi, fattispecie incriminatrici⁵².

Al contrario, secondo altri Autori il mantenimento della sanzione detentiva per la diffamazione, sia pure come *extrema ratio*, non appare inopportuno. Ciò sia in considerazione dell'importanza del bene da tutelare e dei danni irreparabili che la lesione della reputazione può comportare, sia per evitare che gli editori più "ricchi" possano oltrepassare i limiti del lecito, mentre quelli più "poveri" no⁵³. Si è sottolineato che, almeno per i casi più gravi di diffamazione, ovvero le offese seriali, le campagne stampa volte a screditare, le violente offese basate su fatti che il giornalista sa essere false, sia appropriato mantenere come pena la reclusione⁵⁴.

Come anticipato, sul tema è intervenuta la Corte Costituzionale, che è stata chiamata a pronunciarsi circa la compatibilità tra reclusione e diffamazione a mezzo stampa.

La Consulta, riprendendo l'orientamento diffuso all'interno della giurisprudenza di legittimità, ha ritenuto la pena detentiva non incompatibile con il reato di diffamazione, chiarendo che anche talune forme gravi di diffamazione possono trovare come risposta punitiva la detenzione.

Più nel dettaglio, la Corte Costituzionale ha precisato che accanto «ai discorsi d'odio e all'istigazione alla violenza che possono nel caso concreto connotare anche contenuti di carattere diffamatorio»⁵⁵, casi egualmente eccezionali, tali da giustificare l'infrazione di sanzioni detentive, potrebbero «ad esempio essere anche rappresentati da campagne di disinformazione condotte attraverso la stampa, internet o i social media, caratterizzate dalla diffusione di addebiti gravemente lesivi della reputazione della vittima, e compiute nella consapevolezza da parte dei loro autori della – oggettiva e dimostrabile – falsità degli addebiti stessi»⁵⁶. Chi pone in essere simili condotte, continua la Corte, «crea un pericolo per la democrazia, combattendo l'avversario mediante la menzogna, utilizzata come strumento

⁵¹ Cass. pen., sez. V, 26 gennaio 2015, n. 39195.

⁵² In questo senso v. S. Turchetti, *Diffamazione, pena detentiva, caso Sallusti*: cit., e S. Lonati, *Diffamazione a mezzo stampa e applicazione della pena detentiva*, cit.). In quest'ultimo senso A. Gullo, *La tela di Penelope. La riforma della diffamazione nel Testo unificato approvato alla Camera il 24 giugno 2015*, in *Diritto penale contemporaneo*, 1, 2016, 7.

⁵³ Così C. Melzi d'Eril, *La Corte europea condanna l'Italia per sanzione e risarcimento eccessivi in un caso di diffamazione*, cit., 10; Nello stesso senso anche M. Cuniberti, *Pene detentive per la diffamazione, responsabilità del direttore e insindacabilità delle opinioni del parlamentare*, cit., 5 ss.; A. Giudici, *Il caso Belpietro c. Italia: la pena detentiva*, cit.; G.E. Vigevani, *Libertà di espressione, onore e controllo del potere. Sviluppi del diritto di critica politica, tra giudice nazionale ed europeo*, in *Federalismi.it*, 3, 2015, 15.

⁵⁴ Così C. Melzi d'Eril, *La Corte europea condanna l'Italia per sanzione e risarcimento eccessivi in un caso di diffamazione*, cit., 10.

⁵⁵ Corte cost., 12 luglio 2021, n. 150, punto 6.2.

⁵⁶ *Ibid.*

per screditare la sua persona agli occhi della pubblica opinione. Con prevedibili conseguenze distorsive anche rispetto agli esiti delle stesse libere competizioni elettorali»⁵⁷.

La Corte costituzionale ha così fornito una propria interpretazione delle circostanze “eccezionali” che possono giustificare il ricorso alla pena detentiva, individuando ulteriori ipotesi rispetto a quelle enunciate della CEDU.

Nonostante questa precisazione che, di fatto, conferma l’indirizzo già consolidato all’interno della giurisprudenza di legittimità, l’esegesi della cornice entro la quale i giudici potranno continuare ad infliggere la reclusione risulta, come rilevato dalla dottrina più attenta⁵⁸, tutt’altro che agevole.

Dal tenore letterale della pronuncia sembrerebbe che tutte le campagne diffamatorie svolte mediante diffusione di falsità che non implicino alcun rischio per la democrazia non possano essere punite con la reclusione⁵⁹. Cosa debba intendersi per “campagne di disinformazione” e “pericolo per la democrazia” non è tuttavia chiarito.

In ogni caso, deve rilevarsi che, trattandosi di pronuncia interpretativa di rigetto, le indicazioni della Corte Costituzionale esplicano una funzione meramente indicativa e non hanno effetti direttamente vincolanti sui giudici ordinari⁶⁰.

La sentenza in commento, pur riprendendo l’orientamento consolidato, non fornisce indicazioni più precise al riguardo, rimettendo così al giudice di merito il compito di individuare in quali circostanze la condotta diffamatoria possa essere considerata di “eccezionale gravità”.

Quel che è certo è che, per la Cassazione, il ricorso alla sanzione detentiva per il reato di diffamazione, anche se non commesso nell’ambito dell’attività giornalistica, deve essere circoscritto a condotte diffamatorie particolarmente gravi. Del resto se la reclusione è legittima per il giornalista solo in ipotesi eccezionali, lo stesso dovrebbe valere per il “*quisque de populo*” che con l’eventuale condotta diffamatoria difficilmente potrebbe creare un “pericolo per la democrazia”.

6. Conclusioni

La sentenza in commento conferma e avvalorata l’orientamento giurisprudenziale, ormai consolidato, volto a circoscrivere il ricorso alla sanzione detentiva solo in casi di diffamazione grave.

Sicuramente si tratta di una pronuncia da accogliere con favore, sia per il richiamo alla più recente giurisprudenza costituzionale e convenzionale, sia per la sensibilità dimostrata nei confronti di chi non esercita l’attività giornalistica.

Appare, infatti, condivisibile l’estensione dei principi elaborati dalla giuri-

⁵⁷ *Ibid.*

⁵⁸ Così C. Malavenda, *La sentenza n. 150/2021 della Corte Costituzionale*, cit.

⁵⁹ In questo senso C. Magnani, *La Corte costituzionale rivede la detenzione per la diffamazione tramite i media*, cit., 192.

⁶⁰ *Ibid.*

sprudenza convenzionale anche al comune cittadino.

Se il principio merita di essere apprezzato e può rappresentare un “precedente” per future pronunce in tema, la sua portata applicativa rimane incerta. È rimessa, infatti, al giudice di merito una importante opera di interpretazione e di individuazione della “eccezionale gravità” della condotta diffamatoria, sola circostanza che, secondo un’interpretazione costituzionalmente e convenzionalmente orientata, può giustificare l’applicazione della pena detentiva.

Keywords

diffamazione - social network - falsità della notizia - comune cittadino - casi eccezionali.

In tema di responsabilità del *blogger* per il contenuto diffamatorio.

Nota a Cass. civ., sez. III, 27 giugno 2025, n. 17360*

Nicolò Zammarchi

Corte di cassazione, sez. III civ., ord. 27 giugno 2025, n. 17360

Il prestatore di servizi informatici che assuma il ruolo di hosting provider non attivo è, di regola, esente da responsabilità per la pubblicazione di contenuti illeciti immessi da terzi. Tuttavia, una volta che egli acquisisca la consapevolezza della manifesta illiceità dei contenuti – in qualunque modo e non necessariamente a seguito di comunicazione delle autorità competenti – è tenuto ad attivarsi per la loro tempestiva rimozione, pena la perdita dell'esenzione da responsabilità.

Sommario

1. Premessa - 2. Il fatto e lo sviluppo dei procedimenti di merito. - 3. La pronuncia della Cassazione. - 4. La responsabilità del *blogger* per il contenuto diffamatorio *online*: una costante evoluzione. - 5. Qualche considerazione. - 6. Conclusioni

1. Premessa

Con l'ordinanza indicata in epigrafe, la Terza Sezione civile della Cassazione è tornata sull'annosa tematica della responsabilità del *blogger* per l'omessa rimozione del contenuto diffamatorio dal sito *internet*.

La questione non si discosta sensibilmente dagli ultimi arresti interpretativi in materia: essa costituisce anzi un logico e coerente sviluppo della giurisprudenza di legittimità concernente la responsabilità dei prestatori di simili servizi secondo il d.lgs. 70/2003, di attuazione della direttiva 2000/31/CE.

Va subito detto che queste disposizioni sono state abrogate in seguito al recepimento, con d.lgs. 50/2024, della direttiva UE/2018/1808 sui servizi

* L'articolo è stato sottoposto, in conformità al regolamento della Rivista, a referaggio anonimo.

digitali, oggi sostituita dal Reg. UE 2022/2065, c.d. *Digital Services Act*¹.

Tale pronuncia rappresenta nondimeno un valido banco di prova per ricostruire l'evoluzione normativa e giurisprudenziale in un contesto al centro di molteplici sollecitazioni, in massima parte conseguenti agli sviluppi, sempre innovativi, che stimolano nuove forme di interazioni telematiche nella "mediasfera"².

Essa costituisce al contempo un interessante punto di convergenza tra la disciplina civile e quella penale, che induce a riflettere sull'attuale disciplina in materia di responsabilità per il contenuto diffamatorio del gestore del *blog*.

2. Il fatto e lo sviluppo dei procedimenti di merito

L'abbrivio del procedimento è dato da un fatto tanto banale quanto, potenzialmente, sempre più ricorrente: l'utente di un *blog*, ritenendosi diffamato dai commenti scritti da altri utenti in calce a un articolo ivi pubblicato, agisce civilmente per il risarcimento del danno contro il diverso gestore del predetto servizio.

Respinta la domanda in primo grado, l'utente ricorre in appello facendo leva sulla responsabilità del gestore del sito, del quale predica l'omessa attivazione nonostante l'intervenuta conoscenza del contenuto illecito dei commenti che erano stati rilasciati sul *blog*, con conseguente violazione degli artt. 16 e 17 d. lgs. 70/2003³.

¹ In tema, per tutti, S. Braschi, *Il nuovo Regolamento sui servizi digitali: quale futuro per la responsabilità degli Internet Service Provider?*, in *Dir. pen. proc.*, 2023, 367 ss.

² Cfr. R. De Simone, *Presi nella rete. La mente ai tempi del web*, Milano, 2012, 11, che identifica con tale espressione un ambiente ove i *media* elettronici in rete giocano un ruolo fondamentale.

³ Per l'art. 16, concernente l'*hosting*, «1. nella prestazione di un servizio della società dell'informazione, consistente nella memorizzazione di informazioni fornite da un destinatario del servizio, il prestatore non è responsabile delle informazioni memorizzate a richiesta di un destinatario del servizio, a condizione che detto prestatore: a) non sia effettivamente a conoscenza del fatto che l'attività o l'informazione è illecita e, per quanto attiene ad azioni risarcitorie, non sia al corrente di fatti o di circostanze che rendono manifesta l'illiceità dell'attività o dell'informazione; b) non appena a conoscenza di tali fatti, su comunicazione delle autorità competenti, agisca immediatamente per rimuovere le informazioni o per disabilitarne l'accesso. 2. Le disposizioni di cui al comma 1 non si applicano se il destinatario del servizio agisce sotto l'autorità o il controllo del prestatore. 3. L'autorità giudiziaria o quella amministrativa competente può esigere, anche in via d'urgenza, che il prestatore, nell'esercizio delle attività di cui al comma 1, impedisca o ponga fine alle violazioni commesse». Per l'art. 17, inerente all'assenza di un obbligo di sorveglianza, «1. nella prestazione dei servizi di cui agli articoli 14 [*mere conduits*], 15 [*caching*] e 16 [appuntamento, *hosting*], il prestatore non è assoggettato ad un obbligo generale di sorveglianza sulle informazioni che trasmette o memorizza, né ad un obbligo generale di ricercare attivamente fatti o circostanze che indichino la presenza di attività illecite; 2. Fatte salve le disposizioni di cui agli articoli 14, 15 e 16, il prestatore è comunque tenuto: a) ad informare senza indugio l'autorità giudiziaria o quella amministrativa avente funzioni di vigilanza, qualora sia a conoscenza di presunte attività o informazioni illecite riguardanti un suo destinatario del servizio della società dell'informazione; b) a fornire senza indugio, a richiesta delle autorità competenti, le informazioni in suo possesso che consentano

La Corte territoriale rigetta l'impugnazione, sostenendo che la clausola «su comunicazione delle autorità competenti» prevista all'art. 16 c. 1 lett. *b*) delimita tassativamente il presupposto di conoscenza idoneo a far sorgere l'obbligo di rimozione alle sole segnalazioni provenienti alle autorità: la segnalazione proveniente dal soggetto offeso o da terzi, come in questo caso, non potrebbe generare alcuna responsabilità in capo al gestore del *blog*.

Le ragioni del rigetto si rinvergono in due diversi argomenti.

Da un lato, la Corte d'appello sostiene che la normativa interna avrebbe portata più limitata rispetto alla Direttiva comunitaria: sebbene questa prevedesse effettivamente una responsabilità del prestatore di servizi per l'omessa rimozione delle informazioni memorizzate di carattere illecito fondata sulla mera conoscenza, acquisita anche *aliunde*, del carattere illecito di esse, la norma interna avrebbe circoscritto sensibilmente un simile dovere, riferendo appunto di una "comunicazione delle autorità competenti", ciò che determina una cd. "conoscenza qualificata" come unico presupposto del dovere di attivarsi.

Dall'altro lato, la Corte d'appello rileva che il sottostante illecito penale di diffamazione ha natura istantanea, il che impedisce di affermare la responsabilità di un soggetto diverso da quello che lo commette in virtù di fatti successivi al perfezionamento della relativa fattispecie, giacché tali condotte – di omessa rimozione – interverrebbero appunto in un momento nel quale il reato si è già perfezionato.

A fronte del rigetto, l'utente ricorre in Cassazione, articolando tre motivi di ricorso che orbitano intorno alla dedotta violazione degli artt. 16 e 17 d.lgs. 70/2003.

Il primo e il secondo motivo, tra loro connessi ed essenzialmente inerenti a questioni probatorie, vengono rigettati.

Ben diversa sorte ha invece il terzo motivo, con il quale il ricorrente lamenta *tout court* la «violazione o falsa applicazione degli artt. 16 e 17 d.lgs. n. 70/2003».

3. La pronuncia della Cassazione

Il vizio di legge lamentato nel terzo motivo di ricorso, accolto in Cassazione, attiene al presupposto del dovere di rimozione gravante sul *blogger ex* art. 16 d.lgs. 70/2003.

Il ricorrente evidenzia che la Corte d'appello, come il giudice di primo grado, correla erroneamente un simile obbligo esclusivamente alla segnalazione dell'autorità pubblica: nella prospettiva dell'utente, esso deriva però non dalla segnalazione – formale – dell'autorità, bensì dalla conoscenza

l'identificazione del destinatario dei suoi servizi con cui ha accordi di memorizzazione dei dati, al fine di individuare e prevenire attività illecite; 3. Il prestatore è civilmente responsabile del contenuto di tali servizi nel caso in cui, richiesto dall'autorità giudiziaria o amministrativa avente funzioni di vigilanza, non ha agito prontamente per impedire l'accesso a detto contenuto, ovvero se, avendo avuto conoscenza del carattere illecito o pregiudizievole per un terzo del contenuto di un servizio al quale assicura l'accesso, non ha provveduto ad informarne l'autorità competente».

– effettiva – dell’illiceità dei commenti.

La Cassazione accoglie la censura e annulla con rinvio la pronuncia di secondo grado.

A tal fine, la Corte muove da una sintetica ricognizione della sentenza di secondo grado, per la quale, come visto, il prestatore dei servizi sarebbe gravato dal dovere di rimozione dei commenti diffamatori solo in base alla “conoscenza qualificata” della loro illiceità, conseguente a un’apposita comunicazione proveniente esclusivamente dalle autorità competenti.

La Cassazione rileva l’erroneità di una simile impostazione per «consapevole contrasto» con quanto affermato dalla giurisprudenza di legittimità in sede penale e, segnatamente, da Cass. pen., sez. V, 8 novembre 2019, n. 12546.

Per questa pronuncia, definita quale l’ultimo e più recente approdo interpretativo della giurisprudenza penale in materia di responsabilità del *blogger*, «in tema di diffamazione, il “*blogger*” risponde del delitto nella forma aggravata, ai sensi del comma 3 dell’art. 595 c.p., sotto il profilo dell’offesa arrecata “con qualsiasi altro mezzo di pubblicità”, per gli scritti di carattere denigratorio pubblicati sul proprio sito da terzi quando, venutone a conoscenza, non provveda tempestivamente alla loro rimozione, atteso che tale condotta equivale alla consapevole condivisione del contenuto lesivo dell’altrui reputazione e consente l’ulteriore diffusione dei commenti diffamatori». Peraltro, la Corte precisa che il riferimento alla giurisprudenza penale – dunque, alla natura di reato istantaneo della diffamazione – è solo parzialmente conferente, giacché l’art. 16 del d.lgs. n. 70 del 2003, alla lettera a), prevede un particolare regime di responsabilità risarcitoria, quindi meramente civile, del prestatore di servizi: il che esplicita come tale disposizione costituisca una fattispecie di imputazione della responsabilità civile che ben potrebbe concretamente prescindere dal concorso nella responsabilità penale del terzo.

La Cassazione rileva altresì un contrasto interpretativo con la giurisprudenza civile. Anche in questo settore si è infatti affermato che «nell’ambito dei servizi della società dell’informazione, la responsabilità dello “*hosting provider*”, prevista dall’art. 16 del d.lgs. n. 70 del 2003, sussiste in capo al prestatore dei servizi che non abbia provveduto alla immediata rimozione dei contenuti illeciti, oppure abbia continuato a pubblicarli, quando ricorrano congiuntamente le seguenti condizioni: a) sia a conoscenza legale dell’illecito perpetrato dal destinatario del servizio, per averne avuto notizia dal titolare del diritto leso oppure “*aliunde*”; b) sia ragionevolmente constatabile l’illiceità dell’altrui condotta, onde lo “*hosting provider*” sia in colpa grave per non averla positivamente riscontrata, alla stregua del grado di diligenza che è ragionevole attendersi da un operatore professionale della rete in un determinato momento storico; c) abbia la possibilità di attivarsi utilmente, in quanto reso edotto in modo sufficientemente specifico dei contenuti illecitamente immessi da rimuovere» (così Cass. civ., sez. I, 19 marzo 2019, n. 7708; nello stesso senso, riportata in sentenza, Cass. civ., sez. III, ord., 18 agosto 2023, n. 24818).

La Corte ribadisce inoltre che la giurisprudenza – civile – in materia ha precisato come la stessa lettera a) dell’art. 16 d.lgs. 70/2003 escluda la responsabilità risarcitoria per le «informazioni memorizzate a richiesta di

un destinatario del servizio» laddove il prestatore dello stesso «non sia al corrente di fatti o di circostanze che rendono manifesta l'illiceità dell'attività o dell'informazione».

Quindi, tenuto conto che la regola generale è quella per cui l'*internet provider* non è tenuto a un controllo preventivo sulla liceità delle informazioni e sulle attività svolte dai destinatari del servizio, e che la sua responsabilità è correlata alla mera memorizzazione delle informazioni provenienti dagli stessi – e non, sottolinea il provvedimento in commento, al concorso nel reato istantaneo di diffamazione commesso dagli stessi – tale previsione deve operare anche laddove le circostanze «che rendono manifesta l'illiceità dell'attività o dell'informazione» divengano note solo dopo l'avvenuta memorizzazione dell'informazione resa dal terzo: di talché, non potendo più impedirne la pubblicazione, lo stesso deve operarne la rimozione, anche a prescindere dall'intervento delle autorità.

La Cassazione conclude così affermando che l'*hosting provider* ha l'obbligo di rimuovere le informazioni illecite nel momento stesso in cui egli, in qualunque modo, acquisisca la conoscenza di fatti o circostanze che rendano tale illiceità manifesta.

Logica conseguenza di quanto affermato è che la «comunicazione delle autorità competenti» rappresenta solo una tra le fonti di acquisizione della predetta conoscenza, per quanto eventualmente “qualificata”, dunque dotata di rilievo privilegiato.

La Corte precisa, infine, che l'interpretazione così offerta, oltre che risultare conforme al dato testuale dell'art. 16 d.lgs. 70/2003, «è la più coerente con la sua stessa *ratio*, di recepimento della previsione della Direttiva 2000/31/CE nel diritto interno».

Il passaggio argomentativo è ulteriormente arricchito dalla considerazione che vede nell'interpretazione dell'art. 16 d.lgs. 70/2003 fatta propria dalla giurisprudenza di legittimità «un equilibrato temperamento tra la tutela della posizione del prestatore di servizi (non attivo) nella società dell'informazione, la cui responsabilità si è ritenuto opportuno limitare sul piano normativo, per favorire lo sviluppo e la trasmissione dei servizi e delle idee in tale ambito, e la tutela dei soggetti che possono esserne potenzialmente lesi».

In questa prospettiva, il punto di equilibrio emergente dalla normativa europea e interna, tra loro integrate, si muove lungo i seguenti principi di diritto: «il prestatore di servizi informatici che assuma il ruolo di *hosting provider* non attivo va, di regola, esente dalla responsabilità per la pubblicazione delle eventuali informazioni illecite che provengano dai terzi – e, più specificamente, in relazione alla problematica oggetto della presente controversia – per tutti gli eventuali commenti diffamatori inviati dai terzi, ma, una volta che egli acquisisca la consapevolezza della manifesta illiceità degli stessi (in qualunque modo, anche non necessariamente a seguito di una comunicazione delle autorità competenti, sebbene, in tale ultimo caso, possa essere più agevole percepire il carattere “manifesto” dell'illiceità), è tenuto ad attivarsi per rimuoverli tempestivamente, per continuare a godere dell'esenzione dalla indicata responsabilità».

La difformità tra la decisione sottoposta a suo scrutinio e, al contempo, la necessità di dare continuità a questi indirizzi interpretativi porta all'annul-

lamento con rinvio da parte della Cassazione.

4. La responsabilità del blogger per il contenuto diffamatorio online: una costante evoluzione

La pronuncia annotata riporta il dibattito in tema di responsabilità dell'*internet service provider*, e segnatamente del *blogger*, in un solco già tracciato. La Cassazione interviene infatti per ricondurre a sistema una parziale deviazione interpretativa e, contestualmente, per ribadire la sostanziale contiguità logica della responsabilità penale e civile del gestore del *blog* per i commenti diffamatori ivi intervenuti; d'altronde, l'obbligo risarcitorio tipico della sede civile discende pur sempre da un fatto che ha necessario rilievo penale, *ex art.* 595 c.p., peraltro aggravato *ex art.* 595, c. 3, c.p. dall'essere il fatto avvenuto avvalendosi di un mezzo di pubblicità⁴.

Per comprendere appieno la questione è allora necessario ripercorrere l'evoluzione interpretativa circa la responsabilità del *blogger* per tale reato.

A tal fine, va anzitutto precisata una differenza sostanziale che intercorre tra l'*internet service provider* e l'amministratore del *blog*: mentre il primo fornisce accesso alla rete o predispone altri servizi tecnici, senza creare contenuti, ciò che gli consente di beneficiare del *safe harbour* previsto dagli artt. 14 ss. d.lgs. 70/2003, il secondo mette tipicamente a disposizione dei suoi utenti una piattaforma virtuale nella quale si interagisce, spesso, su argomenti creati proprio da questo.

Tale distinzione riverbera i suoi effetti anche sulle conseguenze per gli illeciti commessi nello spazio virtuale di riferimento, non risultando inizialmente chiara l'eventuale applicazione del d.lgs. 70/2003 – dunque, l'esenzione da responsabilità ivi prevista – anche in favore del *blogger*.

Il quesito trova una prima risposta essenzialmente positiva in quelle pronunce che, esclusa per il gestore del *blog* l'applicabilità dell'art. 57 c.p.⁵, ne riconducono l'attività nell'alveo applicativo del d.lgs. 70/2003: l'eventuale

⁴ *Ex multis*, Cass. pen., sez. V, 25 gennaio 2021, n. 13979; Cass. pen., sez. V, 7 giugno 2019, n. 27875; Cass. pen., sez. V, 10 maggio 2017, n. 50187; Cass. pen., sez. V, 14 novembre 2016, n. 4873. In dottrina, per tutti, V. Pezzella, *La diffamazione*, Torino, 2020.

⁵ Com'è noto, per tale disposizione, «salva la responsabilità dell'autore della pubblicazione e fuori dei casi di concorso, il direttore o il vice-direttore responsabile, il quale omette di esercitare sul contenuto del periodico da lui diretto il controllo necessario ad impedire che col mezzo della pubblicazione siano commessi reati, è punito, a titolo di colpa, se un reato è commesso, con la pena stabilita per tale reato, diminuita in misura non eccedente un terzo»; essa disciplina un autonomo reato omissivo improprio – e colposo – del direttore di giornale, il quale è tenuto a impedire la commissione dell'altrui reato: in dottrina, per tutti, E. Musco, voce *Stampa (dir. pen.)*, in *Enc. dir.*, vol. XLIII, Milano, 1990, 633. Dibattuta è l'applicabilità dell'art. 57 c.p. alla testata giornalistica *online*: dopo un lungo periodo nel quale la prassi negava la compatibilità tra l'art. 57 c.p. e il periodico telematico (cfr. Cass. pen. Sez. V, 28 ottobre 2011, n. 44126), un improvviso *revirement* giurisprudenziale estendeva tale responsabilità anche a simili fonti di informazione (cfr. Cass. pen. Sez. V, 23 ottobre 2018, n. 1275, che muove da Cass. sez. Un., 29 gennaio 2015, n. 31022); in tema, I. Pisa, *La responsabilità del direttore di periodico on-line tra vincoli normativi e discutibili novità giurisprudenziali*, in *Dir. pen. proc.*, 2019, 407 ss.; F. Cecchini, *La responsabilità del direttore di periodico telematico ex art. 57 c.p. tra divieto di analogia, "esigibilità" del controllo e prevedibilità dell'esito giudiziario*, in *archiviopenale.it*, 30 aprile 2019.

responsabilità del *blogger* discende così dalla sua qualificazione come soggetto attivo o passivo, come per l'*internet service provider*⁶.

In effetti, per mitigare l'impunità conseguente alla pedissequa applicazione del d. lgs. 70/2003, la giurisprudenza ha elaborato una distinzione tra *internet service provider* attivo e *internet service provider* passivo, individuando nel primo un soggetto che, intervenendo nella rete – dunque, per il *blogger*, nel “diario virtuale” –, è suscettibile di rimprovero per i relativi contenuti diffamatori⁷: di talché la responsabilità del *blogger* è circoscritta alla attiva partecipazione agli illeciti degli utenti⁸.

Successive pronunce estendono però la responsabilità del *blogger* al mantenimento – dunque, alla omessa rimozione – del contenuto diffamatorio *online* quale concorrente *ex art.* 110 c.p. nel reato di diffamazione⁹. La conoscenza della natura illecita della pubblicazione fonderebbe l'obbligo del *blogger* di rimuovere il contenuto diffamatorio. La mancata esecuzione del dovere di rimozione darebbe origine a una volontaria protrazione – appunto, concorsuale – della diffamazione.

Un'altra interpretazione ricostruisce invece la responsabilità del *blogger* alla stregua di un autonomo titolo di reato di diffamazione. In tale prospettiva, la creazione di un sito viene ricondotta a un comportamento positivo che veicola contenuti, da ciò conseguendo che il *blogger* può rispondere del reato di diffamazione laddove sia informato dalla natura illecita delle pubblicazioni ivi intervenute¹⁰. Peraltro, sebbene questa soluzione trovi un suo precedente nel caso “*Google vs Vividown*”¹¹ e un appiglio normativo nell'art.

⁶ Nella giurisprudenza sovranazionale, cfr. CGUE, C-291/2013, *Papasavvas* (2015).

⁷ La Cassazione rileva infatti che «l'*hosting provider* attivo è il prestatore di servizi della società dell'informazione il quale svolge un'attività che esula da un servizio di ordine meramente tecnico, automatico e passivo, e pone, invece, in essere una condotta attiva, concorrendo con altri nella commissione dell'illecito, onde resta sottratto al regime generale di esenzione di cui all'art. 16 del d. lgs. n. 70 del 2003, dovendo la sua responsabilità civile atteggiarsi secondo le regole comuni»: cfr. Cass. civ., sez. I, 21 febbraio 2019, n. 7708, in questa *Rivista*, 23 luglio 2019, con nota di M. Bassini, *La Cassazione e il simulacro del provider attivo: mala tempora currunt*; in tema v. anche L. Tormen, *La linea dura della Cassazione in materia di responsabilità dell'hosting provider (attivo e passivo)*, in *Nuova giur. civ. comm.*, 2019, 1047; F. Frigerio, *Attivo, anche se inconsapevole. Il Tribunale di Roma sanziona Vimeo e conferma i caratteri della responsabilità dell'hosting provider attivo per violazione del diritto d'autore altrui*, in questa *Rivista*, 2, 2019, 258 ss. Nella giurisprudenza sovranazionale, CGUE, C-324/09, *L'Oréal SA e a. c. eBay International AG e a.* (2012).

⁸ Cfr. D. De Natale, *La responsabilità dei fornitori di informazioni in internet per i casi di diffamazione on line*, in *Riv. trim. dir. pen. ec.*, 3, 2009, 572 s.; I. Salvadori, *I presupposti della responsabilità penale del blogger per gli scritti offensivi pubblicati su un blog da lui gestito*, in *Giur. mer.*, 4, 2007, 1076-1077; in giurisprudenza, Cass. pen., sez. V, 16 luglio 2010, n. 35511; Cass. pen., sez. V, 19 febbraio 2018, n. 16751.

⁹ Cfr. Cass. pen., sez. V, 14 luglio 2016, n. 54946, con nota di R. Carbone, *Responsabilità del blogger: parziale revirement della Cassazione?*, in *Cass. pen.*, 2017, 2787 ss.

¹⁰ È la soluzione accolta da Cass. pen., sez. V, 8 novembre 2018 (dep. 2019), n. 12546, in *dirittopenalecontemporaneo.it*, 17 maggio 2019, con nota di C. Pagella, *La Cassazione sulla responsabilità del blogger per contenuti diffamatori (commenti) pubblicati da terzi*; più in generale, v. i commenti di B. Panattoni, *I riflessi penali del perdurare nel tempo dei contenuti illeciti nel cyberspace*, in *sistemapenale.it*, 21 maggio 2020.

¹¹ Cass. pen., sez. III, 17 dicembre 2013, n. 5107.

17 dir. UE/2019/790¹², essa è contraddetta dalla diversa responsabilità *ex* art. 16 c. 1 lett. b) d.lgs. 70/2003 e, soprattutto, è sprovvista di base legale¹³. I più recenti orientamenti di legittimità riconducono la questione nel paradigma concorsuale.

Per Cass. pen., sez. V, 21 settembre 2022, n. 45680¹⁴, ferma restando la non applicabilità dell'art. 57 c.p. all'amministratore del sito *internet* – risultando tale norma applicabile alle sole testate giornalistiche telematiche e non anche ai diversi mezzi informatici di manifestazione del pensiero –, e in assenza di specifiche disposizioni incriminatrici, l'ascrivibilità del fatto-reato al *blogger* deve essere ricostruita in base alle comuni regole del concorso nel reato¹⁵.

La Corte sottolinea inoltre che, per evitare un'attribuzione di responsabilità per fatto altrui – il commento diffamatorio pubblicato da altri –, fondata sulla mera posizione di amministratore del sito, è necessario ravvisare un coefficiente di compartecipazione nel fatto¹⁶: e tale coefficiente viene ravvisato nella «mancata tempestiva attivazione del gestore del blog nella rimozione di proposizioni denigratorie», che «può costituire indice di adesione volontaria ad esse, consentendone l'ulteriore divulgazione».

La pronuncia sottolinea l'insussistenza di una posizione di garanzia *ex* art. 40 cpv. c.p. in capo all'amministratore di *blog*, sprovvisto di poteri giuridici impeditivi di eventi offensivi di beni altrui in assenza di fonti normative che li conferiscano.

Nulla escluderebbe però di ricondurre il – consapevole – mantenimento del commento diffamatorio nell'alveo concorsuale, quale forma di ulteriore replica della offensività realizzata tramite la pubblicazione dello scritto diffamante: e tale è la soluzione che viene infine accolta.

La mancata tempestiva attivazione del gestore del *blog* nella rimozione di proposizioni denigratorie viene quindi ricondotta a una volontaria adesio-

¹² La disposizione esclude le piattaforme di *file-sharing* dalle esenzioni stabilite dalla direttiva sul commercio elettronico. La responsabilità per la condivisione non autorizzata di materiali protetti da diritto d'autore è comunque esclusa se il gestore del sito dimostra di «aver compiuto i massimi sforzi per ottenere un'autorizzazione» ovvero di «aver compiuto, secondo elevati standard di diligenza professionale di settore, i massimi sforzi per assicurare che non siano disponibili opere e altri materiali specifici per i quali [abbia] ricevuto le informazioni pertinenti e necessarie dai titolari dei diritti». Il *provider* comunque non risponde se prova «di aver agito tempestivamente, dopo aver ricevuto una segnalazione sufficientemente motivata dai titolari dei diritti, per disabilitare l'accesso o rimuovere [dal sito] *web* le opere o altri materiali oggetto di segnalazione e aver compiuto i massimi sforzi per impedirne il caricamento in futuro». Peraltro, a mente dell'art. 2, n. 6, la direttiva si applica al «prestatore di servizi della società dell'informazione il cui scopo principale o uno dei principali scopi è quello di memorizzare e dare accesso al pubblico a grandi quantità di opere protette dal diritto d'autore o altri materiali protetti caricati dai suoi utenti, che il servizio organizza e promuove a scopo di lucro».

¹³ *Amplius*, S. Braschi, *op. ult. cit.*, 172 ss.; v. anche Cass. pen., sez. V, 12 gennaio 2021, n. 7220, in questa *Rivista*, 2, 2021, con nota di S. Vimercati, *La responsabilità del gestore del sito per i contenuti pubblicati da soggetti terzi*.

¹⁴ Cass. pen., sez. V, 21 settembre 2022, n. 45680, in *Giur. it.*, 2023, 1154 ss., con nota di S. Raffaele, *Il concorso omissivo in diffamazione del blogger: è responsabilità colpevole?*.

¹⁵ Ovviamente, il *blogger* autore dello scritto denigratorio risponde direttamente di tale reato.

¹⁶ Cfr. Cass. pen., sez. V, 19 febbraio 2018, n. 16751.

ne alle esse, con l'effetto di consentirne l'ulteriore divulgazione, secondo una forma di responsabilità penale che oscilla tra un concorso – successivo – nel reato – a questo punto, permanente – di diffamazione e l'agevolazione – quantomeno inizialmente, colposa – dell'altrui illecito¹⁷.

5. Qualche considerazione

Ricostruito il contesto nel quale si colloca la pronuncia, è possibile svolgere alcune considerazioni sul suo impatto nella conformazione della responsabilità del *blogger*.

Allo scopo, occorre muovere dalla ricognizione del fatto che l'attuale disciplina concorsuale di matrice giurisprudenziale presenta più ombre che luci¹⁸.

Anzitutto, essa cozza con la natura istantanea del reato di diffamazione e a nulla valgono le formule – ambigue¹⁹ – impiegate dalla giurisprudenza, che richiamano una condotta di “mantenimento” del contenuto diffamatorio. Tale responsabilità integra infatti un concorso di persone susseguente che esula però dagli schemi *ex artt.* 110 ss. c.p.²⁰. Né sembra possibile rimodulare la diffamazione *online* alla stregua di un reato permanente, “per esposizione”²¹: semplicemente, tale prospettiva determinerebbe la dissoluzione degli istituti che presidiano la consumazione del reato²² e la totale perdita di significato del concetto di azione²³, senza contare che la descrizione legale della fattispecie di diffamazione risulta francamente univoca nell'indicare un reato istantaneo²⁴.

Ma quand'anche si ammettesse il concorso di persone – successivo – del *blogger* nella diffamazione commessa dall'utente, resterebbe da dimostrarne il dolo concorsuale.

Il punto è delicato e impone di procedere per gradi. La Cassazione trae questo coefficiente dalla consapevolezza del contenuto diffamatorio del commento in capo al *blogger* che non si attivi per rimuoverlo. Trattasi di

¹⁷ Individua in tale conformazione della responsabilità del *blogger* un'ipotesi di “agevolazione” S. Braschi, *Il nuovo Regolamento sui servizi digitali: quale futuro per la responsabilità degli Internet Service Provider?*, in *Dir. pen. proc.*, 2023, 373.

¹⁸ In generale S. Braschi, *Social media e responsabilità penale dell'Internet Service Provider*, in questa *Rivista*, 3, 2020, 170 ss.; v. anche G. Miceli, *Profili evolutivi della responsabilità in rete: il ruolo degli Internet Service Provider tra prevenzione e repressione*, ivi, 2017, 106 ss.

¹⁹ A. Ingrassia, *Responsabilità penale degli internet service provider: attualità e prospettive*, in *Dir. pen. proc.*, 2017, 1625 ss.; *contra*, in termini possibilisti, R. Bartoli, *Brevi considerazioni sulla responsabilità penale dell'internet service provider*, in *Dir. pen. proc.*, 2013, 606.

²⁰ S. Seminara, *La responsabilità penale degli operatori su internet*, in *Dir. inf.*, 1998, 765 ss.

²¹ Cfr. T. Padovani, *Il momento consumativo nei reati commessi col mezzo della stampa*, in *Riv. it. dir. proc. pen.*, 1971, 800 ss.

²² In tema, S. Braschi, *La consumazione del reato. Fondamenti dogmatici ed esigenze di politica criminale*, Milano, 2020, 259 ss.

²³ Sul punto, S. Seminara, voce *Internet*, cit., 570.

²⁴ Cfr. S. Seminara, *La responsabilità penale*, cit., 765, ove si osserva che «i reati (di condotta) fondati su verbi modali come diffondere, divulgare ecc. si consumano nel momento in cui i contenuti illeciti sono resi accessibili da parte del loro autore».

una conclusione niente affatto scontata: la dottrina ha infatti evidenziato che, laddove «esulano i presupposti di operatività dell'art. 40 cpv. c.p., colui che omette un'azione doverosa potrà assumere il ruolo di *partecipe*, ma non in relazione ad un preteso contributo “materiale” di tipo *omissivo*, bensì in funzione del contributo “morale”, consistente nell'eventuale rafforzamento del proposito criminoso in capo all'autore in senso stretto, che sia determinato dalla promessa di astenersi dal dovuto intervento impeditivo»; laddove invece il preteso garante abbia la mera consapevolezza unilaterale di collaborare all'altrui azione delittuosa, e in assenza dei presupposti *ex art. 40 cpv. c.p.*, è esclusa la possibilità di ravvisare nel comportamento dell'omittente un contributo di tipo morale e ci si dovrà rassegnare a considerare l'omissione penalmente irrilevante²⁵.

I commentatori più recenti sottolineano la difficoltà nell'accertamento – induttivo e da effettuare secondo tutte le emergenze del caso di specie²⁶ – di un tale coefficiente soggettivo, che sorregge una condotta, di fatto, di mera inerzia, dunque priva di un apprezzabile sostrato fisico-materiale²⁷.

Queste difficoltà potrebbero trovare un valido, se non decisivo, superamento nella valorizzazione, quale indicatore del dolo, della segnalazione del contenuto diffamatorio, con richiesta di rimozione dello stesso, pervenuta al gestore del *blog* stesso.

È del resto in questo senso che si muove la pronuncia oggetto di odierno commento laddove – sebbene a fini del risarcimento civile, dunque secondo un'ottica che ammette anche di forme di responsabilità oggettiva – sottolinea che la segnalazione dell'autorità costituisce solo uno dei mezzi di acquisizione della notizia fondante il consapevole dovere di attivarsi che segna il momento in cui il *blogger* accede all'altrui illecito, divenendone corresponsabile.

Ed è qui che si apprezza il rilievo della decisione in esame, che può costituire un utile *trait d'union* tra la responsabilità penale e quella civile per uno stesso fatto diffamatorio.

In definitiva, il profilo soggettivo dell'illecito appare ormai sufficientemente delineato. Permane, tuttavia, un margine di incertezza circa la configurazione oggettiva della condotta ascrivibile al *blogger* – oscillante tra un inammissibile concorso successivo nell'altrui reato, un autonomo titolo di illecito ovvero un modello essenzialmente colposo di agevolazione del fatto altrui – che rappresenta nondimeno il nodo centrale della questione, anche ai fini della responsabilità civile.

6. Conclusioni

Nelle pagine che precedono si è ripercorsa una decisione di indubbio interesse per mettere a fuoco l'attuale statuto della responsabilità del *blogger* per

²⁵ L. Riscato, *Combinazione e interferenza di forme di manifestazione del reato. Contributo ad una teoria delle clausole generali di incriminazione suppletiva*, Milano, 459 ss.; più in generale, I. Leoncini, *Obbligo di attivarsi, obbligo di garanzia e obbligo di sorveglianza*, Torino, 1999.

²⁶ In tema, F.M. Iacoviello, *Processo di parti e prova del dolo*, in *Criminalia*, 2010, 463 ss.

²⁷ Cfr. S. Raffaele, *op. ult. cit.*

commento diffamatorio pubblicato sul “diario virtuale” da questi gestito. Analizzando le cadenze motivate della pronuncia, si è visto che la Cassazione propone un modello essenzialmente unitario – civile e penale – di accertamento del fatto diffamatorio del *blogger*, che risulta fondato sulla mancata, consapevole rimozione di un simile commento in conseguenza di una segnalazione anche non qualificata che gli pervenga.

Quanto precede induce peraltro a circoscrivere il rilievo delle segnalazioni che possono fondare una responsabilità in capo al *blogger* unicamente a quelle che siano state effettuate secondo i canali di comunicazione propri del *blog* stesso.

Questa impostazione risponde, da un lato, all’esigenza di legalità e tassatività tipica della sede penale, e, dall’altro, alla necessità di garantire al *blogger* la possibilità concreta di acquisire conoscenza effettiva del contenuto segnalato. Solo le comunicazioni trasmesse attraverso queste modalità, e corredate delle informazioni necessarie a circoscrivere con precisione la natura diffamatoria del messaggio, dovrebbero dunque rilevare ai fini della configurazione del dovere di attivarsi in capo al gestore del *blog*: segnalazioni generiche, formulate in modo vago o inviate tramite vie non ufficiali, non dovrebbero invece essere considerate idonee a far sorgere una posizione di garanzia o un obbligo giuridico in capo al *blogger*.

Sotto altro profilo, l’attuale configurazione della responsabilità penale del *blogger* per l’omessa eliminazione di contenuti diffamatori pubblicati – *recte*, mantenuti – sul proprio sito presenta ancora molteplici criticità, risultando fondata su un modello concorsuale difficilmente sostenibile.

Le incertezze qui compendiate potrebbero trovare soluzione negli sviluppi dell’implementazione del nuovo *Digital Services Act*. Superando il principio della direttiva CE/2000/31, che limitava la responsabilità del fornitore del servizio ai soli casi di dolo diretto e attività non meramente tecniche, il Regolamento modifica i criteri della responsabilità dei *provider*, stabilendo (art. 16, par. 3) che la semplice ricezione di una segnalazione conforme ai requisiti dettagliatamente indicati all’art. 16, par. 1 e 2²⁸, fa scattare la “conoscenza o consapevolezza effettiva” dell’illiceità del contenuto, obbligandolo così a rimuoverlo.

Questo assetto apre alla possibilità di una responsabilità penale di tipo colposo per il *provider* – e, *mutatis mutandis*, del *blogger* – che non elimini i contenuti illeciti segnalati. Certo, anche questa soluzione non sarebbe esente da riserve critiche: oltre a recare con sé inevitabili questioni di proporzionalità

²⁸ Secondo le due disposizioni introdotte dal Regolamento, i prestatori di servizi di memorizzazione di informazioni predispongono meccanismi – di facile accesso e digitali – per consentire a qualsiasi persona o ente di notificare loro la presenza, nel loro servizio, di informazioni che costituiscano contenuti illegali. Tali segnalazioni devono essere sufficientemente motivate; devono indicare l’ubicazione elettronica esatta delle informazioni che si presume costituiscano contenuti illegali; devono indicare il nome e l’indirizzo di posta elettronica del segnalante, salvo riguardino reati di abuso sessuale, sfruttamento sessuale, pornografia minorile e adescamento di minori indicati nella direttiva 2011/93/UE; una dichiarazione con la quale la persona o l’ente conferma la propria convinzione, in buona fede, circa l’esattezza e la completezza delle informazioni e le dichiarazioni ivi contenute. È da notare che il nuovo *Digital Services Act* introduce altresì (art. 22) dei segnalatori attendibili le cui segnalazioni devono essere trattate in via prioritaria. Entrambe le disposizioni costituiscono elementi di novità rispetto all’assetto previgente.

sanzionatoria derivanti dal raffronto con la disciplina *ex art.* 57 c.p. – allo stato non è prevista alcuna riduzione di pena simile a quella, invece, che spetta al direttore di giornale – essa potrebbe determinare l'attribuzione di una responsabilità da posizione, fondata sulla semplice carica di gestore del *blog*²⁹.

Resta allora da vedere in che modo la prassi – e, soprattutto, il futuro legislatore, inevitabilmente chiamato a intervenire sul tema – si adatterà alla nuova realtà normativa.

Keywords

diffamazione - blog - hosting provider - conoscenza dell'illiceità - comunicazione delle autorità competenti.

²⁹ Più nel dettaglio, S. Braschi, *Il nuovo regolamento*, cit., 373 ss.

L'indebito nella gestione collettiva dei diritti d'autore: limiti dell'autotutela economica e responsabilità dell'intermediario*

Edoardo Salvatore

Tribunale di Roma, sez. II civile, 26 ottobre 2025, n. 14852

La SLAE non può procedere unilateralmente alla compensazione di somme indebitamente versate allo Stato mediante decurtazione dei proventi spettanti agli aventi diritto, dovendo invece agire nei confronti del destinatario del pagamento non dovuto. Il diritto alla ripetizione dell'indebito, soggetto a prescrizione decennale, decorre dal riconoscimento giudiziale della titolarità dell'opera e può essere interrotto da diffide stragiudiziali idonee a manifestare la volontà creditoria.

Sommario

1. Premessa. – 2. I fatti di causa e la decisione. – 3. La gestione collettiva dei diritti d'autore: evoluzione normativa e quadro sistematico. – 4. Il delicato confine tra legittimo esercizio di funzioni amministrative e illegittima autotutela economica. – 5. Prescrizione dell'azione di ripetizione e tutela dell'erede. – 6. La perdurante centralità della buona fede e del divieto di ingiustificato arricchimento nella gestione collettiva dei diritti d'autore: correttezza sostanziale e tutela effettiva degli autori.

1. Premessa

La pronuncia del Tribunale di Roma, sez. II civile, 26 ottobre 2025, n. 14852, in commento affronta con rigore e sensibilità sistematica il tema della qualificazione e della ripartizione della responsabilità nella gestione collettiva dei diritti d'autore, allorché l'intermediario – la SIAE – agisca in modo difforme dai principi di correttezza e buona fede nella restituzione di somme indebitamente versate. Il caso offre lo spunto per riflettere sul rapporto, spesso problematico, tra funzione pubblicistica e dimensione

* L'articolo è stato sottoposto, in conformità al regolamento della Rivista, a referaggio anonimo.

contrattuale della gestione collettiva dei diritti autorali, in un contesto in cui la società di intermediazione continua a operare al crocevia tra obbligazioni civili, tutela patrimoniale e interesse generale alla circolazione delle opere dell'ingegno.

2. I fatti di causa e la decisione

La vicenda trae origine dai versamenti che la SIAE effettuò tra il 1974 e il 1981 in favore del Ministero dell'Economia e delle Finanze a titolo di "diritti demaniali" relativi a opere musicali ritenute all'epoca cadute in pubblico dominio. A seguito di successivi accertamenti giudiziari – dapprima con una sentenza del Tribunale di Torino del 2002 e poi con la conferma della Corte di Cassazione nel 2008 – veniva riconosciuta la coautorialità delle opere e, di conseguenza, la perdurante tutela del diritto d'autore in capo agli aventi causa dell'autore originario.

Nonostante tale riconoscimento, le somme corrisposte al Ministero erano già state interamente versate. In luogo di promuovere un'azione di ripetizione dell'indebito nei confronti del destinatario del pagamento, la società di gestione collettiva attraverso varie operazioni contabili successive provvedeva a detrarre dai proventi maturati e spettanti agli eredi dell'autore la quota corrispondente a quanto indebitamente versato allo Stato, giustificando l'operazione come recupero compensativo di somme non dovute. Gli aventi diritto convenivano in giudizio la SIAE e il Ministero chiedendo la restituzione delle somme trattenute e, in via subordinata, la condanna solidale dei convenuti al pagamento delle rispettive quote. Il Ministero eccepiva la prescrizione decennale del diritto alla ripetizione, sostenendo che il termine decorreva dalla data del versamento originario; la SIAE, dal canto suo, contestava la propria legittimazione passiva e invocava la buona fede e la correttezza del proprio operato, richiamando una scrittura del 2008 contenente una rinuncia o comunque un assenso all'operazione di recupero. In via subordinata, la società di gestione formulava domanda riconvenzionale nei confronti del Ministero per essere rimborsata delle somme eventualmente dovute agli eredi.

Il Tribunale ha ricostruito le diverse posizioni distinguendo il profilo della prescrizione da quello della legittimità dell'operato della SIAE. In primo luogo, ha accolto l'eccezione del Ministero, dichiarando prescritto il diritto alla ripetizione nei suoi confronti, in quanto l'azione era stata proposta oltre il termine decennale decorrente dal momento in cui il diritto poteva essere esercitato. Diversamente, ha ritenuto non prescritta la pretesa nei confronti della SIAE, valorizzando la continuità di atti interruttivi del termine, rappresentati da diffide e richieste di rendiconto inviate dagli eredi negli anni successivi e, in particolare, da una diffida del marzo 2018.

Nel merito, il giudice del Tribunale di Roma ha ritenuto illegittima la detrazione operata dalla SIAE sui proventi spettanti agli aventi diritto, affermando che l'intermediario non poteva surrogarsi al debitore originario né trasferire sui titolari dei diritti il rischio economico derivante da un pagamento indebito. La società avrebbe dovuto richiedere la restituzione direttamente al Ministero in quanto destinatario della prestazione non do-

vuta e non agire mediante compensazione unilaterale.

Per tali motivi, il Tribunale adito ha accolto la domanda degli eredi nei confronti della SIAE condannandola al pagamento della quota di spettanza e, contestualmente, ha accolto la domanda riconvenzionale della stessa società di gestione collettiva nei confronti del Ministero, condannando quest'ultimo a rimborsare la medesima somma.

3. La gestione collettiva dei diritti d'autore: evoluzione normativa e quadro sistematico

La pronuncia in esame non può essere pienamente compresa se non alla luce del quadro normativo e istituzionale entro cui la gestione collettiva dei diritti d'autore ha progressivamente assunto forma e rilievo. È opportuno, perciò, soffermarsi, sia pure per cenni, sull'evoluzione del sistema, sulle ragioni del monopolio legale della SIAE e sulle trasformazioni introdotte a livello europeo e interno negli ultimi anni.

La gestione collettiva dei diritti d'autore e dei diritti connessi trova la propria giustificazione nell'esigenza di assicurare un più efficiente coordinamento delle attività di negoziazione e di controllo relative allo sfruttamento economico delle opere dell'ingegno¹. Tale obiettivo viene perseguito attraverso la centralizzazione di tali funzioni in capo a soggetti intermediari qualificati, in grado di garantire una tutela effettiva e uniforme degli interessi degli autori e degli altri titolari dei diritti.

La necessità di un sistema di gestione collettiva si è manifestata in modo particolarmente evidente con riferimento alle opere musicali² – quali composizioni, canzoni e produzioni audiovisive – la cui diffusione capillare e la molteplicità delle modalità di utilizzo hanno reso difficoltosa, se non del tutto impraticabile, la concessione di licenze basate su accordi individualmente negoziati dai titolari dei diritti. In questi ambiti, la gestione accentrata si è, pertanto, imposta quale strumento indispensabile per garantire un equilibrio tra l'effettiva fruibilità delle opere da parte del pubblico e la necessaria tutela degli interessi economici degli autori e dei produttori³.

Nel corso del tempo, a livello europeo, si è così consolidato un modello di gestione accentrata, caratterizzato da un marcato elemento di territorialità

¹ In argomento, *ex multis*, si vedano: S. Izzo, *Entità di gestione indipendenti e sistema italiano d'intermediazione dei diritti d'autore: riflessioni sulla sentenza Lea c. Jamendo*, in *Dir. un. eur.*, 3, 2024, 559 ss.; A. Giannaccari, *Prove di liberalizzazione del mercato (italiano) della gestione collettiva dei diritti d'autore: buona la seconda?*, in *il Foro it.*, 9, 2024, 445 ss.; G. Remotti, *Studio sulle dinamiche contrattomercato nella gestione collettiva dei diritti d'autore e connessi*, in *Contr. e impr.*, 2, 2024, 442 ss.; C. Meo, *Gestione collettiva e mercato dei diritti d'autore*, Torino, 2023; P. Dodaro, *I diritti connessi e il ruolo degli enti intermediari*, in E. Battelli (a cura di), *Diritto privato dello spettacolo*, Torino, 2020, 61 ss.; M. Ficsor, *La gestion collective du droit d'auteur et des droits connexes*, Genève, 2002; R. Mastroianni, *I diritti esclusivi delle collecting societies*, in *Annali it. dir. autore*, 2001, 202 ss.; Id., *Diritto internazionale e diritto d'autore*, Milano, 1997; L.C. Ubertazzi, *Diritto d'autore e internazionale e comunitario*, in *Dig. disc. priv. Sez. comm.*, Torino, 1989, 350 ss.

² Cfr. E. Battelli – E.M. Incutti, *Le opere musicali*, in E. Battelli (a cura di), *Diritto privato dello spettacolo*, cit., 135 ss. e, in specie, 141 ss. e 146 ss.

³ Per un'analisi economica del fenomeno, si veda M. Lamandini, *Televisioni, nuove tecnologie e IP: nuovi mercati e collecting societies*, in *AIDA*, 2010, 135 ss.

e fondato sull'articolazione dei mercati nazionali. In tale contesto, specifici organismi – le cosiddette *collecting societies*⁴ – sono stati investiti del compito di negoziare, concedere e vigilare sull'utilizzazione economica delle opere musicali tutelate, fungendo da intermediari istituzionali tra i titolari dei diritti e gli utilizzatori⁵. Si è dunque progressivamente affermata una tendenza al consolidamento di situazioni monopolistiche – talora previste *ex lege*, talaltra di fatto – fondata sull'assunto che la presenza di un unico soggetto intermediario per ciascun ordinamento nazionale potesse rappresentare un elemento di efficienza e di equilibrio nel sistema⁶. Tale assetto, infatti, è stato ritenuto idoneo, da un lato, ad accrescere la forza contrattuale nella negoziazione delle condizioni di sfruttamento delle opere a vantaggio degli autori e, dall'altro, a garantire agli utilizzatori una sensibile semplificazione delle procedure di rilascio delle licenze. Ebbene, nell'ordinamento giuridico italiano fino al 2017 la SIAE⁷ ha go-

⁴ P. Dodaro, *I diritti connessi e il ruolo degli enti intermediari*, cit., 98 ss.

⁵ In forza del mandato conferito dai titolari dei diritti d'autore, le *collecting societies* rilasciano licenze agli utilizzatori delle opere, provvedendo alla riscossione dei compensi dovuti e alla loro successiva ripartizione tra i titolari, previa detrazione di una quota a copertura dei costi connessi all'attività di intermediazione svolta. Tra le funzioni loro attribuite rientra, altresì, quella di prevenire e contrastare ogni forma di utilizzazione non autorizzata delle opere protette, assicurando così un'effettiva tutela dei diritti economici e morali degli autori.

⁶ Rispetto a tale configurazione di carattere eminentemente nazionale e tendenzialmente monopolistica, l'Unione europea ha adottato, nel corso del tempo, un orientamento evolutivo. In una fase iniziale, infatti, il settore è stato ricondotto esclusivamente nell'ambito di applicazione delle norme del diritto dell'Unione in materia di concorrenza, con particolare riferimento ai divieti relativi all'abuso di posizione dominante. Solo in un momento successivo, tuttavia, si è avvertita l'esigenza di introdurre una disciplina specifica della gestione collettiva, volta a colmare le criticità che l'esperienza applicativa aveva dimostrato non potersi risolvere attraverso la sola dinamica concorrenziale del mercato. Cfr. G.M. Riccio – G.G. Codiglione, *Copyright Collecting Societies Monopolistic Positions and Competition in the EU Single Market*, in *Masaryk University Journal of Law and Technology*, 7, 2013, 1 ss.; F. Siirainen, "Droit d'auteur" contra "droit de la concurrence": versus "droit de la regulation", in *Rev. int. droit économique*, 4, 2001, 413 ss. Invero, la svolta è stata determinata, in misura rilevante, dai progressi tecnologici i quali, rendendo possibile un controllo a distanza sulle modalità di sfruttamento dei diritti di proprietà intellettuale, hanno progressivamente attenuato – se non del tutto eliminato – la necessità di mantenere una rigida connotazione territoriale dei sistemi di gestione collettiva. Tale superamento ha contribuito a rimuovere uno dei principali ostacoli alla piena liberalizzazione del mercato in questo ambito. Ne è conseguita, dunque, l'adozione di interventi normativi da parte dell'Unione europea volti ad armonizzare la disciplina dei diritti d'autore e dei diritti connessi, in una prospettiva di maggiore integrazione e concorrenzialità del settore. Al riguardo A.M. Rovati, *La gestione collettiva dei diritti d'autore: brevi note sulla liberalizzazione del settore*, in *Ist. Dir. Eco. [on line]*, 3, 2019, 4 ss.

⁷ Fondata nel 1882 su iniziativa di esponenti del mondo artistico e intellettuale, la Società Italiana degli Autori (poi Società Italiana degli Autori ed Editori – SIAE, dal 1927) nacque come sodalizio privato finalizzato alla tutela collettiva dei diritti d'autore. In origine, l'adesione poteva avvenire in forma associativa o mediante mandato di rappresentanza, a seconda che l'autore intendesse far parte o meno dell'organizzazione. Nel corso del tempo, la SIAE ha assunto una connotazione progressivamente pubblicistica, dapprima per effetto dell'inquadramento corporativo del periodo fascista, poi con la legge sul diritto d'autore del 1941 che le attribuì funzioni e prerogative di carattere pubblico, tra le quali proprio la gestione esclusiva di determinati diritti (art. 180 l.d.a.) e la tenuta di registri ufficiali. Ad oggi, l'art. 1 della Legge del 9 gennaio 2008, n. 2, qualifica la SIAE come ente pubblico economico a base associativa, nel quale convivono due profili distinti: da un lato, la funzione di interesse generale di vigilanza e promozione culturale; dall'altro, la gestione economica dei diritti degli autori e mandanti secondo principi di efficienza, trasparenza e concorrenzialità. Sul ruolo e sulle

duto di un monopolio legale esclusivo nell'attività di intermediazione, in qualunque forma esercitata, per la gestione dei diritti espressamente individuati dall'art. 180 l.d.a., tra i quali i diritti di rappresentazione, esecuzione, recitazione, radiodiffusione e riproduzione meccanica e cinematografica. Con riferimento a tali diritti, l'ente era l'unico soggetto legittimato sul territorio nazionale a svolgere in forma collettiva le attività di rilascio delle licenze agli utilizzatori, riscossione dei proventi e successiva ripartizione dei compensi tra gli aventi diritto.

Per lungo tempo, l'esclusiva legale riconosciuta alla SIAE è stata ritenuta conforme a Costituzione. La stessa Corte Costituzionale ne individuava la *ratio* giustificativa nelle funzioni di interesse pubblico svolte dall'ente volte a promuovere la cultura e la diffusione delle opere dell'ingegno, nonché a garantire un'efficace tutela dei diritti d'autore⁸. Si riteneva, infatti, che l'attività di intermediazione, monitoraggio e riscossione dei diritti richiedesse una complessa struttura organizzativa diffusa sull'intero territorio nazionale, i cui costi potessero essere sostenuti solo in presenza di un regime di monopolio legale⁹.

Analogo modello, peraltro, si riscontrava in numerosi altri Paesi europei in cui la gestione collettiva dei diritti d'autore e connessi era tradizionalmente affidata a soggetti nazionali operanti in condizioni di monopolio o oligopolio¹⁰. La riserva dell'attività di intermediazione ad un unico ente per ciascun territorio favoriva l'aggregazione dei repertori su base nazionale e comportava che il rilascio di licenze per utilizzazioni transfrontaliere – ad esempio, lo sfruttamento di un'opera britannica in Italia o viceversa – avvenisse mediante accordi bilaterali di rappresentanza. Tali accordi, promossi e regolati nell'ambito del sistema confederativo CISAC, prevedevano che ciascuna *collecting society* nazionale affidasse la gestione del proprio repertorio all'estero alle società omologhe, ricevendo da queste un'analoga delega per il territorio di competenza. Si veniva così a consolidare un sistema fondato su una sostanziale reciprocità ed esclusività nella gestione dei repertori¹¹.

I progressi tecnologici¹² – e, in particolare, lo sviluppo delle tecnologie di

funzioni svolte dalla SIAE, si rinvia a G. Jarach – A. Pojaghi, *Manuale del diritto d'autore*, Milano, 290 ss.; nonché a A. Sirotti Gaudenzi, *Il nuovo diritto d'autore*, Santarcangelo di Romagna, 2024, 158 ss.

⁸ Si veda, ad esempio, Corte cost., 15 maggio 1990, n. 241, in *Dir. aut.*, 365 ss.

⁹ Si riteneva, pertanto, che soltanto un regime di monopolio legale potesse garantire l'efficienza economica del sistema, consentendo l'attivazione di un meccanismo di compensazione dei costi capace di beneficiare al contempo delle “economie di scala” – derivanti dalla concentrazione a livello nazionale delle risorse in capo ad un unico ente – e delle “economie di scopo”, ossia dall'impiego diversificato delle medesime risorse per lo svolgimento di una pluralità di attività di vigilanza e di controllo sulle varie forme di utilizzazione delle opere dell'ingegno.

¹⁰ Interessante è lo studio di D. Menegon, *L'intermediazione dei diritti d'autore*, in *IBL Briefing papers*, 19 luglio 2010, 89 ss., in cui si evidenzia che fino al 2010 vi era un unico monopolio legale in Danimarca, Austria, Belgio, Repubblica Ceca ed Italia. Tuttavia, già nel 2016 il quadro risultava profondamente mutato dal momento che solo la Repubblica Ceca conservava un monopolio legale nell'intermediazione dei diritti d'autore.

¹¹ L. Marcheggiani, *Le licenze multi-territoriali per l'uso online di opere musicali nella disciplina comunitaria della gestione collettiva dei diritti d'autore*, in *Oss. Dir. Civ. e Comm.*, 2, 2013, 293 ss.

¹² Con riferimento alla tendenza alla “disintermediazione” derivante dall'avvento delle nuove

gitali – hanno profondamente trasformato i meccanismi di intermediazione dei diritti d'autore, ampliando le modalità di sfruttamento delle opere e favorendo la nascita di forme alternative di gestione¹³. Tale evoluzione, da un lato, ha esteso il novero dei diritti di privativa suscettibili di tutela e, dall'altro, ha accresciuto le esigenze di protezione e valorizzazione da parte di autori, artisti e interpreti. Parallelamente, essa si è intrecciata con un più ampio processo di liberalizzazione che, secondo una dinamica comune ad altri settori economici, ha finito per investire anche l'ambito della gestione collettiva dei diritti d'autore e dei diritti connessi.

In tale scenario, caratterizzato dal fallimento prodotto da tale modello monopolistico-nazionalistico di «gestione collettiva a vocazione universale»¹⁴, sono dunque apparsi diversi e ulteriori soggetti intermediari, tra i quali spiccano le c.d. «entità di gestione indipendenti» (EGI) e gli «organismi di gestione collettiva» (OCG), entrambi previsti dalla direttiva 2014/26/UE del Parlamento europeo e del Consiglio del 26 febbraio 2014 – la c.d. “Direttiva Barnier” – concernente la gestione collettiva dei diritti d'autore e dei diritti connessi, nonché la concessione di licenze multi territoriali per i diritti sulle opere musicali destinate all'uso online nel mercato interno¹⁵. Tale intervento normativo è stato adottato con l'obiettivo di armonizzare le discipline nazionali in materia di accesso e funzionamento delle *collecting societies*, al fine di rendere più trasparente, efficiente e concorrenziale l'atti-

tecnologie, si vedano A.M. Gambino – A. Stazi, *Diritto dell'informatica e della comunicazione*, Torino, 2009, 1 ss.; nonché M. Ricolfi, *Figure tecniche di gestione del diritto d'autore e dei diritti connessi*, in P. Spada (a cura di), *Gestione collettiva dell'offerta e della domanda di prodotti culturali*, Milano, 2009, 6 ss., il quale fornisce una visione prospettica del nuovo ruolo istituzionale affidato alle *collecting societies*. È stato osservato come l'attività di produzione professionale esercitata nell'ambito del diritto d'autore, e in particolare nella distribuzione dei contenuti digitali, abbia in realtà sostituito al carattere dell'intermediazione quello dell'“organizzazione e dell'aggregazione”. Al riguardo si veda: E. Prosperetti, *L'opera digitale tra regole e mercato*, Torino, 2017.

¹³ Ciascun autore è infatti in grado di amministrare e disporre individualmente dei propri diritti. Si pensi, ad esempio, alle c.d. “*Licenze Creative Commons*” (M. Kim, *The Creative Commons and Copyright Protection in the Digital Era: Uses of Creative Commons Licenses*, in *Journal of Computer-Mediated Communication*, 2007, 187 ss.), o più recentemente alla possibilità di caricare le proprie opere sulle principali piattaforme di *streaming* musicale (cfr. E.M. Incutti – F. Ruggeri, *La circolazione dei diritti tramite internet: il ruolo delle piattaforme digitali*, in E. Battelli (a cura di), *Diritto privato dello spettacolo*, cit., 286 ss.).

¹⁴ Per «gestione collettiva a vocazione universale» si intende quel modello organizzativo in cui la *collecting society* mira a rappresentare la totalità dei titolari dei diritti, indipendentemente dalla loro dimensione economica o dalla specifica tipologia dei diritti amministrati. Si tratta di un modello volto a soddisfare interessi «solidaristici a favore di soggetti dotati di minore forza contrattuale, quali tipicamente sono gli autori». In questi termini, P.M. Sanfilippo, *La gestione collettiva dei diritti d'autore e connessi tra regolazione e concorrenza*, in *AIDA*, 2007, 445 ss.

¹⁵ In altri termini, la direttiva 2014/26/UE, in ragione del nesso indissolubile che lega il mercato e il contratto, mira a disciplinare i rapporti negoziali a monte e a valle al fine di conformare in modo equilibrato entrambe le dimensioni del mercato della gestione collettiva dei diritti. Sul tema, A. Belvedere, *Le dinamiche contratto-mercato*, in Aa.Vv., *Contratto e mercato. Liber Amicorum per Angelo Luminoso*, Milano, 2013, 237 ss.; R. Natoli, *Il contratto “adeguato”*, Milano, 2012, 43 ss.; A. Nervi, *Il contratto come strumento di conformazione dell'assetto di mercato*, in *Eur. dir. priv.*, 1, 2018, 102 ss. Con particolare riferimento alla distinzione tra “disciplina normativa del mercato” e “disciplina del contratto” si rinvia a N. Irti, *L'ordine giuridico del mercato*, Bari, 1998, 42 ss.; nonché V. Roppo, *Il contratto*, in *Trattato di diritto privato*, a cura di G. Iudica e P. Zatti, Milano, 2011, 871 ss.

vità di gestione collettiva¹⁶.

Il d.lgs. n. 35 del 15 marzo 2017, pur recependo la Direttiva Barnier, ne attuò solo parzialmente le previsioni, disciplinando l'accesso e l'organizzazione degli OGC e delle EGI, lasciando tuttavia impregiudicata la riserva d'intermediazione a favore della SIAE prevista dall'art. 180 l.d.a. Ne derivò, dunque, un quadro di sostanziale ambiguità: formalmente riconosciuta la libertà di affidamento a organismi anche stranieri, questi ultimi non potevano operare direttamente in Italia poiché dovevano comunque avvalersi della SIAE quale intermediario necessario.

Solo con il d.l. n. 148 del 16 ottobre 2017 (conv. in l. n. 172/2017) il legislatore ha definitivamente rimosso il monopolio legale, consentendo anche ad altri organismi di gestione collettiva – ma non anche alle entità di gestione indipendenti – di svolgere direttamente attività di intermediazione. Tale facoltà resta subordinata alla condizione che tali organismi siano controllati o partecipati dai titolari dei diritti o costituiti in forma associativa senza fini di lucro, in funzione di garanzia dell'interesse collettivo degli autori.

Sebbene la riforma del 2017 abbia formalmente liberalizzato il mercato della gestione collettiva dei diritti di cui all'art. 180 l.d.a., la SIAE conserva tuttora un ruolo centrale, in virtù di numerose prerogative esercitate in via esclusiva. Le restano, infatti, funzioni di riscossione e ripartizione di compensi derivanti da diritti diversi, come l'equo compenso per reprografia (artt. 68 e 181-*ter* l.d.a.), per copia privata (artt. 71-*septies* e 71-*octies* l.d.a.)¹⁷ e per il diritto di seguito (artt. 144 ss. l.d.a.)¹⁸. Ulteriori competenze riguardano la determinazione dei compensi per l'esecuzione pubblica delle opere radiodiffuse (art. 58 l.d.a.) e la gestione delle comunioni ereditarie coatte (art. 116, c. 2, l.d.a.). Accanto a tali attività privatistiche, la SIAE continua a svolgere funzioni di rilievo pubblicistico, quali l'accertamento e la riscossione di imposte e diritti (art. 181 l.d.a.), la prevenzione delle violazioni (artt. 182-*bis* e *ter* l.d.a.)¹⁹, il rilascio di contrassegni antipirateria (art. 181-*bis* l.d.a.) e la tenuta di registri pubblici, tra cui quello dei programmi per elaboratore (art. 103, c. 4, l.d.a.)²⁰.

¹⁶ Nel mutato contesto normativo europeo, l'Autorità Garante della Concorrenza e del Mercato (AGCM) aveva sollecitato l'Italia ad adeguare il proprio ordinamento ai principi della direttiva 2014/26/UE, superando il monopolio legale riconosciuto alla SIAE. Tale assetto, infatti, appariva non più compatibile con l'evoluzione tecnologica e con l'esigenza di garantire ai titolari dei diritti un'effettiva libertà di scelta tra gli intermediari collettivi. Cfr. Parere AGCM AS1281 del 1° giugno 2016, richiamato anche da A. Sirotti Gaudenzi, *Il nuovo diritto d'autore*, cit., 138.

¹⁷ Il meccanismo di ripartizione dei compensi si articola in tre fasi principali: la SIAE raccoglie i proventi dovuti, ne trasferisce una quota alle associazioni di categoria più rappresentative tramite specifici accordi e queste ultime provvedono, infine, a distribuirli proporzionalmente tra i propri associati e mandanti.

¹⁸ A differenza dell'intermediazione dei diritti di cui all'art. 180 l.d.a., la riscossione dell'equo compenso e del diritto di seguito da parte della SIAE non deriva da un mandato conferito dai titolari dei diritti, ma trova fondamento direttamente nella legge.

¹⁹ Gli artt. 182-*bis* e 182-*ter* l.d.a. attribuiscono ai funzionari SIAE, in collaborazione con le unità AGCOM, funzioni ispettive assimilabili a quelle di pubblici ufficiali, conferendo pubblica fede alle loro attestazioni.

²⁰ P. Dodaro, *I diritti connessi e il ruolo degli enti intermediari*, cit., 106.

Ne emerge, in definitiva, la persistente natura “ibrida” dell’ente, sospesa tra funzioni pubblicistiche di tutela collettiva e compiti di gestione economica nell’interesse dei titolari dei diritti.

4. Il delicato confine tra legittimo esercizio di funzioni amministrative e illegittima autotutela economica

La posizione peculiare che la SIAE occupa nel sistema – al tempo stesso soggetto di rilevanza pubblica e gestore privato di interessi collettivi – rende centrale la questione dei limiti entro i quali essa possa esercitare poteri di gestione e di controllo economico sulle somme amministrate. Proprio tale dimensione ibrida, che ne legittima il ruolo di intermediario necessario, impone tuttavia un vincolo rafforzato di conformità ai principi di correttezza e buona fede, i quali sono destinati a permeare ogni fase dell’attività gestoria.

Il caso deciso dal Tribunale di Roma evidenzia in modo emblematico le tensioni derivanti da tale duplicità di natura. Dopo aver riconosciuto l’errore nel versamento di somme al Ministero dell’Economia e delle Finanze, la SIAE ha proceduto autonomamente ad un recupero contabile, trattando parte dei compensi spettanti agli aventi diritto. Questa operazione, qualificata come compensazione interna, è stata giustificata dall’Ente come misura di riequilibrio patrimoniale, ma proprio tale condotta rivela la delicatezza del confine tra legittimo esercizio di funzioni amministrative e illegittima autotutela economica.

Come noto, l’ordinamento civile riconosce solo in via eccezionale il potere del creditore di soddisfare unilateralmente le proprie ragioni, subordinandolo ad un fondamento normativo espresso o a un rapporto di reciprocità. Nel caso di specie, nessuna di queste condizioni ricorreva: gli aventi diritto non erano debitori della SIAE per il pagamento indebito effettuato in favore del Ministero, né sussisteva un titolo che legittimasse la detrazione. Il Tribunale ha, dunque, correttamente escluso la possibilità di ricondurre tale condotta ad una forma di compensazione legittima, ravvisandovi, invece, un’indebita ingerenza nel patrimonio dei rappresentati.

L’esigenza di impedire simili forme di autotutela trova il fondamento nei principi di correttezza e buona fede che, in un rapporto fiduciario come quello di gestione collettiva, assumono una funzione di preservazione di un equilibrio negoziale strutturale. L’art. 1175 c.c., infatti, impone al debitore e al creditore di comportarsi secondo lealtà nell’adempimento delle obbligazioni, mentre l’art. 1375 c.c., a sua volta, estende tale dovere alla fase di esecuzione del contratto, fungendo da criterio di integrazione del contenuto obbligatorio²¹. In applicazione di tali principi, la SIAE non

²¹ Sulla nozione di buona fede in senso oggettivo e in senso soggettivo, *ex multis*, si richiamano gli studi di: C. Grassetti, *L’interpretazione del negozio giuridico con particolare riguardo ai contratti*, Padova, 1938, 191 ss.; L. Barassi, *La teoria generale delle obbligazioni. III. L’attuazione*, Milano, 1946, 795 ss.; R. Sacco, *La buona fede nella teoria dei fatti giuridici di diritto privato*, Torino, 6 ss.; F. Carusi, *Correttezza (obblighi di)*, in *Enc. dir.*, X, Milano, 1962, 709 ss.; G. Giampiccolo, *La buona fede in senso soggettivo nel diritto privato*, in *Riv. dir. comm.*, 1965, I, 335 ss.; A. Cataudella, *Sul*

può perseguire finalità di tutela patrimoniale proprie mediante strumenti che pregiudichino gli autori rappresentati, perché tale agire avrebbe come conseguenza quello di alterare il sinallagma fiduciario che giustifica il suo ruolo.

La pronuncia in commento riafferma, in questa prospettiva, un principio di portata generale: la gestione collettiva dei diritti d'autore, pur svolgendo funzioni di rilievo pubblico, rimane soggetta ai canoni civilistici, i quali escludono ogni forma di autotutela patrimoniale priva di base legale. Ne discende che la SIAE, nell'amministrare i proventi spettanti agli autori, deve attenersi ad un modello di condotta improntato a trasparenza, correttezza e parità di trattamento, coerentemente con la natura fiduciaria del mandato ricevuto. Solo così può realizzarsi quel bilanciamento tra efficienza gestionale e tutela individuale che rappresenta, da sempre, la giustificazione stessa della gestione collettiva.

5. Prescrizione dell'azione di ripetizione e tutela dell'erede

Accanto al profilo dell'indebito e dei limiti dell'autotutela economica, la decisione del Tribunale affronta anche la questione, non meno rilevante, della prescrizione del diritto alla ripetizione e della sua incidenza sulla posizione giuridica degli eredi dell'autore. L'aspetto è particolarmente significativo, poiché il giudice è chiamato a bilanciare l'esigenza di certezza dei rapporti giuridici con quella di assicurare una tutela effettiva a soggetti che non avevano preso parte diretta alle vicende originarie del rapporto di gestione.

Il Tribunale ha distinto con chiarezza le due relazioni obbligatorie coinvolte: da un lato quella intercorrente tra la SIAE e il Ministero dell'Economia e delle Finanze e, dall'altro, quella tra la SIAE e gli aventi diritto. Per quanto concerne il rapporto con l'Amministrazione, il giudice ha accolto l'eccezione di prescrizione sollevata dal Ministero, ritenendo che il termine decennale dell'azione di ripetizione fosse ampiamente decorso. La causa di estinzione del diritto è stata ricondotta al momento in cui la SIAE avrebbe potuto esercitare l'azione di ripetizione verso il destinatario del pagamento indebito, ossia sin dalla definitività dell'erronea qualificazione delle opere come appartenenti al pubblico dominio.

Diversa è la valutazione compiuta in ordine al rapporto tra la SIAE e gli eredi dell'autore. In questo caso, il Tribunale ha ritenuto che l'azione di ripetizione proposta dagli aventi diritto non fosse prescritta, riconoscendo efficacia interruttiva ad una serie di atti stragiudiziali – diffide, richieste di rendiconto, istanze di trasparenza – intervenuti negli anni successivi. In

contenuto del contratto, Milano, 1966, 254; S. Rodotà, *Le fonti di integrazione del contratto*, Milano, 1969, 132 ss.; D. Corradini, *Il criterio della buona fede e la scienza del diritto privato. Dal codice napoleonico al codice civile italiano del 1942*, Milano, 1970, 137 ss.; G. Alpa, *Pretese del creditore e normativa di correttezza*, in *Riv. dir. comm.*, 1971, II, 286; F.D. Busnelli, *Buona fede in senso soggettivo e responsabilità per fatto "ingiusto"*, in *Studi sulla buona fede*, Milano, 1975, 569 ss.; M. Bessone – A. D'Angelo, *Buona fede*, in *Enc. giur. Treccani*, V, Roma, 1988, 1 ss.; nonché da ultimo F. Piraino, *La buona fede in senso oggettivo*, Torino, 2015.

particolare, la diffida del marzo 2018 è stata considerata manifestazione inequivoca della volontà creditoria, idonea a interrompere il decorso della prescrizione ai sensi dell'art. 2943, c. 4, c.c.²².

La decisione si fonda su una ricostruzione che valorizza la continuità del rapporto di gestione e la posizione di dipendenza informativa in cui versano gli autori e i loro eredi rispetto alla SIAE. In un contesto nel quale i titolari dei diritti non dispongono autonomamente di tutti gli elementi contabili necessari per verificare l'esattezza delle ripartizioni. L'applicazione rigorosa delle regole in materia di prescrizione, pertanto, deve essere temperata dal principio di buona fede che impone al gestore di cooperare alla realizzazione dell'interesse del creditore. È in questa prospettiva che la giurisprudenza tende a riconoscere efficacia interruttiva a diffide e solleciti, anche reiterati, con cui gli aventi diritto manifestino la volontà di ottenere chiarimenti o rettifiche sulle operazioni contabili effettuate dall'intermediario.

L'impostazione seguita dal Tribunale consente di evitare che la prescrizione diventi strumento di ingiustificato arricchimento per il gestore, così riaffermando un principio di equità sostanziale nei rapporti di durata, nei quali il decorso del tempo non può essere valutato in modo avulso dal comportamento della parte che ne trae beneficio. L'erede dell'autore, in quanto successore nella titolarità dei diritti e nella posizione contrattuale di gestione, subentra in un rapporto già caratterizzato da asimmetrie informative e deve poter contare sulla lealtà dell'intermediario nel garantire la piena trasparenza delle operazioni pregresse.

In tal senso, la sentenza valorizza la funzione equilibratrice e di garanzia dell'art. 2935 c.c., riaffermando che il termine di prescrizione decorre solo da quando il titolare è in condizione di esercitare il proprio diritto²³. Nei rapporti di gestione collettiva, tale condizione non può dirsi realizzata fino a quando il gestore non renda conoscibili gli elementi essenziali per la verifica dei conteggi e delle imputazioni contabili. È solo in presenza di tale trasparenza che può considerarsi pienamente tutelata l'esigenza di certezza giuridica che la prescrizione è destinata a soddisfare.

²² In generale, sul tema del decorso della prescrizione si rinvia, tra gli altri, a: G. Azzariti – G. Scarpello, *Della prescrizione e della decadenza. Artt. 2900-2969*, VI, *Della tutela dei diritti*, in *Commentario del codice civile* a cura di A. Scialoja – G. Branca, Bologna-Roma, 1953, 640 ss.; A. Auricchio, *Appunti sulla prescrizione*, Napoli, 1971, 56 ss.; B. Grasso, voce «Prescrizione (dir. priv.)», in *Enc. dir.*, XXXV, Milano, 1986, 57 ss.; P. Vitucci, *Le disposizioni generali sulla prescrizione*, in F. Roselli – P. Vitucci, *La prescrizione e la decadenza*, in *Trattato di diritto privato*, diretto da P. Rescigno, vol. 20, *Tutela dei diritti*, t. II, Torino, 1998, 468 ss.; M. Tescaro, *Decorrenza della prescrizione e autoreponsabilità. La rilevanza civilistica del principio contra non valentem agere non currit praescriptio*, Padova, 2006, 77 ss.; P.G. Monateri, *La prescrizione*, in P.G. Monateri – C. Costantini, *La parte generale del diritto civile*, in *Trattato di diritto civile*, diretto da R. Sacco, Torino, 2009, 28 ss.

²³ Per tutti, C.M. Bianca, *Le garanzie reali. La prescrizione*, in *Diritto civile*, 7, Milano, 2012, 555 ss.

6. La perdurante centralità della buona fede e del divieto di ingiustificato arricchimento nella gestione collettiva dei diritti d'autore: correttezza sostanziale e tutela effettiva degli autori

La sentenza del Tribunale di Roma in commento si inserisce in una linea interpretativa che, pur muovendo da un caso specifico, contribuisce a chiarire alcuni nodi strutturali del sistema della gestione collettiva dei diritti d'autore. Essa riafferma la perdurante centralità dei principi civilistici – in particolare della buona fede, della correttezza e del divieto di ingiustificato arricchimento – come parametri capaci di orientare l'attività di un soggetto che, sebbene investito di funzioni di rilievo pubblicistico, resta vincolato alle regole dell'autonomia privata e della responsabilità contrattuale.

In tale prospettiva, la decisione appare significativa perché ridimensiona la pretesa autosufficienza organizzativa della SIAE, riconducendone l'operato nell'alveo del diritto comune delle obbligazioni. L'intermediario collettivo non può sottrarsi al principio di causalità che governa le attribuzioni patrimoniali, né sostituire al rimedio della *repetitio indebiti*, ricorrendo a strumenti di compensazione interna che ne alterino la *ratio*. L'errore di imputazione economica, anche quando dipenda da esigenze gestionali, non giustifica l'imposizione di un sacrificio patrimoniale agli autori rappresentati, la cui posizione resta tutelata dal vincolo fiduciario che fonda l'intera architettura del sistema.

Senza assumere tratti di rottura, la pronuncia segna un passo di consolidamento nella direzione di una civilizzazione della gestione collettiva, intesa come riconduzione delle sue dinamiche entro i confini della responsabilità ordinaria. Essa dimostra che l'efficienza del modello non richiede l'ampliamento di spazi di autotutela, ma piuttosto la valorizzazione delle regole di lealtà e trasparenza che garantiscono l'affidamento degli autori e la legittimazione dell'ente stesso.

In questa luce, la vicenda giudiziaria appare come un richiamo alla responsabilità sistemica della SIAE e, più in generale, degli organismi di gestione collettiva nel conformare la propria azione a criteri di correttezza sostanziale e di tutela effettiva degli autori.

In definitiva, l'equilibrio tra funzione collettiva e interesse individuale – che costituisce la cifra identitaria del modello – può conservarsi solo se il gestore continua a riconoscere, anche nei casi più complessi, la priorità del principio di giustizia commutativa su ogni logica di convenienza economica.

Keywords

diritto d'autore - gestione collettiva dei diritti d'autore - autotutela economica - buona fede e correttezza - indebito oggettivo

Cronache

La stretta su intercettazioni e ordinanze cautelari: la stampa, da cane da guardia a sciacallo che si nutre di pettegolezzo?*

Carlo Melzi d'Eril

*Non posso concepire che una mano pura
possa toccare un giornale senza fremere di disgusto*
Charles Baudelaire

Sommario

1. Premessa: le disposizioni originarie e gli interessi in gioco. - 2. I correttivi alla pubblicabilità delle intercettazioni telefoniche. - 3. L'inedito divieto di pubblicazione delle ordinanze di custodia cautelare.

1. Premessa: le disposizioni originarie e gli interessi in gioco

Le novità legislative in tema di pubblicazione di intercettazioni telefoniche (l. 114/2024¹) e di ordinanze di custodia cautelare (d.lgs.

* L'articolo è stato sottoposto, in conformità al regolamento della Rivista, a referaggio anonimo.

¹ L. 114/2024, *Modifiche al codice penale, al codice di procedura penale, all'ordinamento giudiziario e al codice dell'ordinamento militare*, in *Gazz. Uff.*, 10 agosto 2024, n. 187. Per un quadro e alcuni primi commenti sulle disposizioni di natura processuale introdotte dalla c.d. riforma Nordio, dal disegno di legge al testo definitivo della legge, si rinvia a M. Gialuz, *Le novità della "manovra Nordio" in materia processuale: quando l'ideologia rischia di provocare un'eterogeneità dei fini*, in *Sistema penale*, 22 luglio 2024; G. Illuminati, *Le modifiche al processo penale nel d.d.l. Nordio: una prima lettura*, in *Rivista italiana di diritto e procedura penale*, 2023, 883 ss.; L. Ludovici, *Disegno di legge c.d. Nordio: nuove garanzie processuali tra fughe in avanti e false partenze*, in *lalegislazionepenale.eu*, 7 maggio 2024; E. Maccora, *Il cantiere sempre aperto della giustizia penale. Primi osservazioni al DDL Nordio ed all'impatto sulle sezioni gip-gup*, in *questionegiustizia.it*, 6 luglio 2023; F. Porcu, *Le modifiche al codice di rito nel d.d.l. Nordio: un primo commento alla riforma in fieri*, in *Penale. Diritto e Procedura*, 2 agosto 2023; G. Spangher, *Pacchetto Nordio: timidi ma significativi segnali di cambio di prospettiva*, in *Penale. Diritto e Procedura*, 27 giugno 2023; E. Valentini, *Le novità della "legge Nordio" in materia cautelare: poche idee, ma confuse*, in *lalegislazionepenale.eu*, 17 febbraio 2025. E, più nello specifico, sulle interpolazioni in materia di intercettazioni v. A. Apollonio, *D.d.l. Nordio in materia di intercettazioni: l'ennesima*

198/2024²) mostrano impietosamente – se si ha una qualche dimestichezza con la materia – come l'attuale governo consideri il rapporto fra processo penale e informazione e, in particolare, quale ruolo attribuisca alla stampa.

Inutile cercare risposte nelle dichiarazioni di questo o quello schieramento politico, troppo spesso illogiche, inconcludenti, ridotte a meri *slogan* per scopi elettorali.

Uno strumento più credibile a disposizione dell'interprete è l'analisi degli interventi normativi e delle loro conseguenze, con particolare attenzione agli equilibri di volta in volta stabiliti a favore di questo o quell'interesse coinvolto. Ad esempio, rivelatore è il maggiore privilegio garantito alla libertà di diffondere notizie di pubblico interesse oppure, viceversa, di aumentare il perimetro del segreto garantito agli atti di indagine.

Per comprendere la *ratio* delle norme oggetto di riforma è indispensabile analizzare la disciplina originaria in materia di pubblicazione degli atti del processo penale contenuta nel codice di rito³. La materia ha subito, nel tempo, stratificazioni e aggiustamenti. Stando però all'architettura generale della normativa, si può dire che i beni giuridici in discussione, in origine almeno, erano essenzialmente due: da un lato, il corretto e regolare svolgimento del processo penale che postula l'efficacia dell'attività investigativa,

ombra gettata sull'operato del pubblico ministero (e l'ennesimo passo verso la separazione delle carriere), in giustiziainsieme.it, 27 maggio 2024; F. Demartis, Intercettazioni: un correttivo che rafforza la riservatezza delle comunicazioni e tutela l'attività del difensore... ma scontenta i "curiosi", in Diritto penale e processo, 2024, 1168 ss.; L. Filippi, Intercettazioni: un intervento legislativo all'insegna della tutela della privacy, in Le novità della legge Nordio, a cura di G. Spangher, Milano, 2024, 21 ss.; W. Nocerino, Le recenti riforme in materia di intercettazioni: alcune criticità e qualche prospettiva, in lalegislazionepenale.eu, 17 febbraio 2025.

² D. lgs. 198/2024, Disposizioni per il compiuto adeguamento della normativa nazionale alle disposizioni della direttiva (UE) 2016/343 del Parlamento europeo e del Consiglio, del 9 marzo 2016, sul rafforzamento di alcuni aspetti della presunzione di innocenza e del diritto di presenziare al processo nei procedimenti penali, in *Gazz. Uff.*, 23 dicembre 2024, n. 300. Per un quadro e alcuni primi commenti v., tra gli altri, G. Amato, Un intervento normativo condivisibile che vuole superare prassi non corrette, in *Guida al diritto*, 2, 2025, 80; G. Giostra, Vietare la pubblicazione dell'ordinanza di custodia cautelare: una proposta funzionalmente e sistematicamente in difficoltà di senso, in *Sistema penale*, 7 ottobre 2024; A. Zampini, Divieto di pubblicazione delle ordinanze che applicano misure cautelari personali: una scelta politica censurabile e costituzionalmente repressibile, in *questionegiustizia.it*, 13 dicembre 2024.

³ Sulla genesi dei divieti di pubblicazione originariamente inseriti nel codice di rito del 1989 e sui beni giuridici tutelati nella prospettiva del legislatore codicistico, v., per tutti, G. Giostra, *Processo penale e informazione*, Milano, 1989, 279 ss., ed E. Lupo, sub art. 114, in *Commento al nuovo codice di procedura penale*, II, coord. da M. Chiavario, Torino, 1990, 43 ss. Più in generale, sui divieti di pubblicazione – senza pretesa di completezza – *ex multis*: R. I. Adorno, *Sulla pubblicazione del contenuto degli atti di indagine coperti dal segreto*, in *Cassazione penale*, 1995, 2167; R. Bartoli, *Tutela penale del segreto processuale e informazione: per un controllo democratico sul potere giudiziario*, in *Diritto penale contemporaneo*, 3, 2017, 59; L. Carli, *Indagini preliminari e segreto investigativo*, in *Rivista italiana di diritto e procedura penale*, 1994, 76; M. Ceresa Gastaldo, *Processo penale e cronaca giudiziaria: costituzionalmente illegittimo il divieto di pubblicazione degli atti del fascicolo per il dibattimento*, in *Giurisprudenza costituzionale*, 1995, 2123; G. Giostra, *Alt al divieto di pubblicare gli atti del fascicolo del dibattimento*, in *Diritto penale e processo*, 1995, 659; F. Palazzo, *Note sintetiche sul rapporto tra giustizia penale e informazione giudiziaria*, in *Diritto penale contemporaneo*, 3, 2017, 141; N. Triggiani, *Giustizia penale e informazione. La pubblicazione di notizie, atti e immagini*, Padova, 2012.

a “verginità” cognitiva e la serenità di giudizio del magistrato⁴; dall’altro, la libertà di espressione, e quale suo necessario corollario, il diritto della collettività a essere informata (anche) sull’operato dell’Autorità giudiziaria, così da poter esercitare un controllo tempestivo ed effettivo sul potere giudiziario e, in primo luogo, sull’azione della Pubblica accusa.⁵

Gli atti di indagine sono coperti dal segreto fino a quando l’imputato non ne può avere conoscenza (art. 329, c. 1, c.p.p.) e per ovvie ragioni non sono pubblicabili *tout court*. Una volta caduto il segreto, se si procede al dibattimento, il contenuto di quanto conservato nel fascicolo formato dagli inquirenti diventa ostensibile, pur restando tuttavia vietata la diffusione del “virgolettato” fino alla pronuncia di secondo grado.

La regola è un corollario del principio del “doppio fascicolo”, che limita l’osmosi delle “carte” tra indagini e dibattimento. Più precisamente, tale divieto è volto a garantire la terzietà e l’imparzialità del giudice, affinché egli non venga a conoscenza, per mezzo della stampa, di elementi esclusi

⁴ Tali temi sono ampiamente trattati da G. Giostra, *Processo penale e informazione*, cit., spec. 299 ss.

⁵ Nella *Relazione al codice di procedura penale*, in *Gazz. Uff.*, 24 ottobre 1988, n. 250, Suppl. ord. n. 2, 177, si sottolineava la necessità di un «tempestivo controllo della pubblica opinione e degli organi di informazione sul comportamento tenuto dalla pubblica accusa proprio in un sistema, come quello nuovo, in cui la piena autonomia in cui questa viene a trovarsi elimina (o quantomeno attenua) qualsiasi altro tipo di controllo su eventuali omissioni». In dottrina, evidenzia il compito fondamentale dell’informazione nel controllo sull’operato del potere giudiziario G. Giostra, *Processo penale e informazione*, cit., 288, osservando che la nuova fisionomia delle indagini preliminari ha persino «accresciuto l’importanza del controllo della pubblica opinione in correlazione alla diminuita incidenza del controllo endoprocessuale esercitabile dalla difesa dell’inquisito, che è di regola estromessa da questa fase a carattere marcatamente preprocessuale e di parte». In termini analoghi, R. Bartoli, *Tutela penale del segreto processuale e informazione: per un controllo democratico sul potere giudiziario*, cit., 74 afferma che «in una democrazia nessun potere incontrollabile è ammissibile. (...) Anche il potere giudiziario indipendente deve essere riportato entro la dimensione di un controllo che non può che essere democratico, in modo che – senza alcun pregiudizio per l’indipendenza della giustizia – anche la giustizia resti espressione reale, non di un ceto burocratico separato e irresponsabile, ma della sovranità popolare. Se così stanno le cose, ecco allora che la funzione dell’informazione pubblica non è soltanto quella di fornire informazioni affinché la collettività formi una propria opinione sui fatti, ma diviene anche l’istanza di controllo popolare dinanzi alla quale il potere giudiziario indipendente può essere chiamato a rendere conto». Sulla medesima linea già D. Pulitanò, “Potere d’informazione” e “giustizia”: per un controllo democratico sulle istituzioni, in N. Lipari (a cura di), *Giustizia e informazione*, Roma-Bari, 1975, 143 ss. Più in generale, sul ruolo della stampa quale – per utilizzare il linguaggio della giurisprudenza sovranazionale – “cane di guardia della democrazia” (*public-watchdog role of the press*), ossia quale soggetto chiamato a vigilare sull’operato dei poteri pubblici, incluso quello giudiziario, quasi in funzione di un “quarto potere”, occorre rinviare anzitutto alla giurisprudenza della Corte EDU, che ha coniato l’espressione e ne ha progressivamente precisato la portata. Si vedano, tra le altre: CEDU, *Bladet Tromsø e Stensaas c. Norvegia*, ric. 21980/93 (1999), §§ 59 e 62; CEDU, *Pedersen e Baadsgaard c. Danimarca*, ric. n. 49017/99 (2004), § 71; CEDU, *Von Hannover c. Germania* (n. 2), ric. 40660/08 e 60641/08 (2012), § 102; CEDU, *Goodwin c. Regno Unito*, ric. 17488/90 (1996), § 39.

Sul medesimo tema – e per ulteriori richiami giurisprudenziali – v. anche la *Guida sull’articolo 10 della Convenzione europea dei diritti dell’uomo. Libertà di espressione* (aggiornata al 30 aprile 2021), ove si ribadisce che «la Corte ha sempre sottolineato il ruolo fondamentale di “cane da guardia” che svolge la stampa in una società democratica, e ha collegato la funzione dei giornalisti di diffondere informazioni e idee su tutte le questioni di interesse generale al diritto, per il pubblico, di riceverne». La Guida è reperibile all’URL giustizia.it.

dal materiale utilizzabili per la decisione⁶.

La regola che pone il segreto sulle indagini, con la peculiare *ratio* e la limitatezza temporale che la contraddistingue, consente di ritenere ragionevole, anzi forse doveroso, un divieto radicale di pubblicazione. Il principio del doppio fascicolo, invece, si mostra più blando per molte ragioni e non sembra consentire un simile radicale oscuramento. La distinzione fra contenuto pubblicabile e virgolettato non pubblicabile, quindi, ha rappresentato il punto di caduta del bilanciamento dei vari interessi in gioco. Un punto di caduta, oggi possiamo dirlo, mal localizzato.

Nel 1989, come oggi, non si sarebbe potuto sacrificare del tutto il diritto della collettività a esercitare uno scrutinio sull'attività del potere giudiziario anche, e forse a maggior ragione, nella fase cruciale delle indagini. Del resto, se «*la giustizia è amministrata in nome del popolo*» (art. 101 c. 1 Cost.), non si può precludere del tutto proprio al popolo di conoscere quel che accade dentro e i tribunali (o le procure) per poterlo valutare, criticare e in fin dei conti controllare⁷. Dunque, da queste premesse si spiega la distinzione contenuta nell'art. 114 c.p.p. tra gli atti coperti dal segreto e quelli non più soggetti a tale vincolo.

Nel contesto odierno, esistono nuove sensibilità e, con queste, sono emersi altri beni giuridici che reclamano attenzione, mentre all'epoca della redazione del codice furono poco considerati, almeno in questo ambito.

A tal proposito, anche per impulso del legislatore sovranazionale e della relativa giurisprudenza, si è consolidata l'idea che il divieto di pubblicazione avrebbe tra gli altri scopi quello di garantire il diritto alla riservatezza⁸.

⁶ Sul tema, v. G. Giostra, *Processo penale e informazione*, cit., spec. 299 ss.; A. Nappi, *Contro le virgolette: il garantismo delle perifrasi*, in *questionegiustizia.it*, 19 marzo 2024; R. Casiraghi, *Informazione giudiziaria, processo mediatico e imparzialità del giudice*, in *Archivio penale*, 3, 2021, 1 ss.

⁷ È pacifico che tra l'art. 21 Cost. e l'art. 101 Cost. insista uno stretto nesso, poiché il loro combinato disposto costituisce presupposto irrinunciabile per l'attuazione del principio democratico. In questo senso v. G. Giostra, *La giustizia penale nello specchio deformante della cronaca giudiziaria*, in questa *Rivista*, 3, 2018, 23, ove l'autore parla di un «filo quasi trasparente ma robustissimo che collega l'art. 21 all'art. 101»: «dal loro combinato disposto, per usare una trita ma difficilmente sostituibile locuzione, si evince che il popolo ha diritto di sapere come viene amministrata la giustizia in suo nome e come i giudici, soggetti soltanto alla legge, esercitano il loro potere di *jus dicer*». Negli stessi termini, Id., *Segreto processuale, tra esigenze di giustizia, diritti individuali e controllo democratico*, in *Cassazione penale*, 2018, 747 ss. Cfr. anche F. Palazzo, *Note sintetiche sul rapporto tra giustizia penale e informazione giudiziaria*, cit., 140, ove si osserva che «il collegamento col principio democratico non sembra che possa avallare una “mediazione” della stampa nel senso di un coinvolgimento emotivo del pubblico nella celebrazione dei singoli processi. Piuttosto, quel principio pone prima di tutto l'esigenza di un controllo democratico sull'amministrazione della giustizia e, in secondo luogo, l'esigenza di una partecipazione popolare alla politica della giustizia attraverso le forme della rappresentanza democratica: ma è evidente che la partecipazione alla politica della giustizia non è realizzabile senza un'accurata e puntuale informazione sulla realtà dell'amministrazione della giustizia».

⁸ La Racc. (2003) 13, adottata il 10 luglio 2003 dal Consiglio d'Europa e contenente *Principi relativi alle informazioni fornite attraverso i mezzi di comunicazione in rapporto a procedimenti penali* dedicava specifico spazio alla «Tutela della privacy in rapporto a procedimenti penali in corso», prevedendo che «nel fornire informazioni relative a persone sospettate, imputate o condannate oppure ad altri soggetti coinvolti in procedimenti penali si dovrebbe rispettare il diritto di tali persone alla tutela della privacy, conformemente

nonché il principio di non colpevolezza della persona sottoposta alle indagini⁹. Ciò, specie in un contesto mediatico ove l'indagato o l'imputato è spesso "vittima" di una semplificazione che lo vede schiacciato su tesi giornalistiche non di rado colpevoliste, essendo come ovvio gli inquirenti le principali fonti della stampa, per lo meno all'inizio della vicenda giudiziaria, se non altro perché solo costoro posseggono le notizie.

Il modo in cui il legislatore ha dato spazio alla tutela di questi beni giuridici è la questione che sarà trattata in questa sede. Lo strumento utilizzato per perseguire un simile risultato è stato quello di estendere – forse in modo eccessivo, sicuramente in maniera maldestra – la latitudine del segreto.

2. I correttivi alla pubblicabilità delle intercettazioni telefoniche

L'istituto delle intercettazioni telefoniche è tanto delicato quanto peculiare. Delicato, poiché si tratta di un mezzo di ricerca della prova molto efficace ma che si scontra, comprimendolo, con (almeno)¹⁰ un diritto fondamentale

all'articolo 8 della Convenzione». Con la sentenza CEDU, *Drakešas c. Lithuania*, ric. 36662/04 (2012), la Corte europea ha espressamente riconosciuto che la pubblicazione di intercettazioni da parte dei media può, in determinate condizioni, integrare una violazione dell'art. 8 CEDU, dedicato alla tutela della riservatezza. In dottrina è oggi generalmente riconosciuto che la riservatezza costituisca un valore fondamentale da proteggere, tanto con riguardo ai soggetti direttamente coinvolti quanto – e soprattutto – ai terzi. Tale esigenza rappresenta, in effetti, uno dei principali motori propulsori delle più recenti riforme in materia di divieti di pubblicazione e, in particolare, di selezione e diffusione delle intercettazioni. Sul punto, ex multis, v. F. Demartis, *Intercettazioni: un correttivo che rafforza la riservatezza delle comunicazioni e tutela l'attività del difensore... ma scontenta i "curiosi"*, cit.; W. Nocerino, *Le recenti riforme in materia di intercettazioni: alcune criticità e qualche prospettiva*, cit.; A. Pugliese, *Pubblicazione di atti processuali e tutela della riservatezza*, in *Diritto di Difesa*, n. 3, 2022, 517 ss.

⁹ Ne è prova il d.lgs. 198/2024, con il quale il legislatore italiano ha adeguato la normativa nazionale alle disposizioni della direttiva 2016/343 del Parlamento europeo e del Consiglio del 9 marzo 2016, rafforzando il principio di presunzione di innocenza dell'imputato e dell'indagato – nella sua accezione costituzionale e sovranazionale – e intervenendo sulla previsione dell'art. 114 c.p.p. Sul rapporto tra pubblicità del processo – che, a causa delle distorsioni dell'informazione, può tradursi in un "processo mediatico", parallelo al processo celebrato in aula – e presunzione di non colpevolezza, la produzione scientifica è sterminata. A titolo meramente esemplificativo, e senza alcuna pretesa di completezza, si rinvia a: G. Caneschi, *Processo penale mediatico e presunzione di innocenza: verso un'estensione della garanzia?*, in *Archivio penale*, 3, 2021, 1 ss.; C. Conti, *Cronaca giudiziaria e processo mediatico: l'etica della responsabilità verso nuovi paradigmi*, in *Archivio penale*, 1, 2022, 1 ss.; G. Giostra, *Processo mediatico*, in *Enc. dir.*, Annali II, t. X, Milano, 2017, 647 ss.; V. Manes, *Giustizia mediatica. Gli effetti perversi sui diritti fondamentali e sul giusto processo*, Napoli, 2022; N. Triggiani, *Introduzione. È la stampa bellezza! E tu non puoi farci niente! Niente! (...neppure con il soccorso della presunzione di innocenza)*, in Aa. Vv., *Informazione e giustizia penale. Dalla cronaca giudiziaria al "processo mediatico"*, Bari, 2022, 1 ss; volendo Aa.Vv., *L'attuazione della direttiva europea sulla presunzione di innocenza*, sezione monografica a cura di C. Melzi d'Eril, in questa *Rivista*, 1, 2022, 11 ss. Per una disamina di alcuni casi emblematici di "processo mediatico" v. Aa. Vv., *Processo mediatico e processo penale. Per un'analisi critica dei casi più discussi da Cogne a Garlasco*, a cura di C. Conti, Milano, 2016.

¹⁰ A seconda dei casi, infatti, oltre alla segretezza delle comunicazioni possono subire lesione ulteriori diritti fondamentali. Cfr. W. Nocerino, *Le recenti riforme in materia di*

tutelato dalla Costituzione, quale quello alla libertà e alla segretezza delle comunicazioni (art. 15 Cost.).

Peculiare, per il modo in cui le operazioni di captazione vengono realizzate e cristallizzati i relativi risultati. In effetti, risuona spesso l'immagine delle intercettazioni come una "pesca a strascico"¹¹, dove nella grande quantità di comunicazioni registrate sarà necessario poi distinguere e separare le informazioni pertinenti e rilevanti rispetto al tema di prova da quelle che non lo sono.

L'obiettivo di selezionare il materiale utile viene perseguito nella cosiddetta "udienza stralcio" (art. 268 c.p.p.), nel corso della quale il giudice acquisisce i flussi di comunicazione di interesse, stralciando appunto il superfluo o quanto coperto dal segreto.

Le affermazioni scambiate in una conversazione privata, quindi, vengono sottratte alla riservatezza loro garantita dalla Costituzione, al fine di tutelare un altro bene giuridico, ovvero l'amministrazione della giustizia. La "ferita" procurata ai diritti della persona non si limita al fatto che i protagonisti del processo verranno a sapere quanto gli interlocutori si sono detti nel (creduto) riserbo. La pubblicità che, al cessare del segreto, caratterizza gli atti del processo fa sì che le parole, spese sotto il rassicurante ombrello della confidenza, diventino di dominio pubblico.

intercettazioni: alcune criticità e qualche prospettiva, cit.: «Diritto di difesa, diritto di cronaca, diritto del singolo e dei consociati al rispetto della privacy, inviolabilità del domicilio, segretezza della corrispondenza, sono solo alcuni dei valori da bilanciare con l'altrettanto essenziale interesse all'accertamento dei reati e della responsabilità penale», cui è funzionale lo strumento delle intercettazioni. Il delicato e precario equilibrio tra i beni giuridici coinvolti rappresenta una delle cause principali delle riforme che si sono succedute negli ultimi anni in materia di intercettazioni. Sul punto, la produzione scientifica è sterminata. Senza alcuna pretesa di completezza, e con riferimento a questo e ad altri profili critici, v., *ex multis*: F. Caprioli, *Intercettazioni e tutela della privacy nella cornice costituzionale*, in *Cassazione penale*, 2021, 1141 ss.; M. Bonetti, *Riservatezza e processo penale*, Milano, 2003, 256 ss.; P. Bruno, voce *Intercettazioni di comunicazioni o conversazioni*, in *Dig. pen.*, X, Torino, 2001, 175; A. Camon, *Le intercettazioni nel processo penale*, Milano, 1996, 44; C. Conti, *La riservatezza delle intercettazioni nella "delega Orlando"*, in *Rivista trimestrale - Diritto penale contemporaneo*, 3, 2017, 78 ss.; L. Filippi, *Intercettazioni telefoniche*, in *Enc. dir.*, VI, Milano, 2002, 565; Id., *L'intercettazione di comunicazioni*, Milano, 1997, 16; A. Gaito – S. Fürfaro, *Intercettazioni: esigenze di accertamento e garanzie della riservatezza*, in Aa. Vv., *I principi europei del processo penale*, a cura di A. Gaito, Roma, 2016, 363 ss.; C. Marinelli, *Intercettazioni processuali e nuovi mezzi di ricerca della prova*, Milano, 2007, 56; A. Nappi, *Sull'utilizzazione extrapenale dei risultati delle intercettazioni*, in *Cassazione penale*, 2014, 386; Id., *Appunti sulla nuova disciplina delle intercettazioni*, in *giustiziainsieme.it*, 17 aprile 2019; N. Triggiani, *Introduzione. È la stampa bellezza! E tu non puoi farci niente! Niente! (...neppure con il soccorso della presunzione di innocenza)*, cit., 1 ss. Per un quadro generale delle riforme susseguitesì prima della riforma Nordio, v. Aa. Vv., *La nuova disciplina delle intercettazioni*, a cura di P. Maggio, Torino, 2023, passim. Sulle prospettive di ulteriore intervento riformatore, v. M. Biral, *Nuovi orizzonti di tutela per i terzi colpiti da intercettazioni*, in *Cassazione penale*, 2025, 1346 ss.

¹¹ La problematica è nota e ampiamente segnalata dalla dottrina, soprattutto a seguito dell'introduzione del c.d. *captatore informatico*. *Ex multis*: F. G. Capitani, *Quali limiti alle "intercettazioni a strascico"?*, in *Diritto e giustizia*, 3, 2020, 5; G. M. Flick, *Intercettazioni, qui si colpisce la base della nostra civiltà. Guai a giocare sui diritti*, in *Cassazione penale*, 2020, 1804; S. Schiavone, *Intercettazioni a "strascico": prospettive interpretative*, in *Giurisprudenza italiana*, 11, 2021, 2487 ss.; G. Spangher – M. Antinucci, *Possibili le intercettazioni "a strascico" attraverso l'uso del captatore informatico per i reati comuni?*, in *Penale. Diritto e Procedura*, 16 marzo 2020, 1 ss.

Sappiamo che ciò è necessario per mantenere trasparenza e controllo su uno dei poteri più incisivi e violenti a disposizione della pubblica amministrazione. Tuttavia, un simile sacrificio sembra accettabile se funzionale al perseguimento dei reati (o al controllo del potere a cui questo compito è affidato), e dunque quando i dati concernono l'imputazione. Inaccettabile appare, invece, la divulgazione di informazioni che, pur magari di straordinario interesse pubblico, riguardano fatti estranei al processo.

In materia di intercettazioni, dunque, la tutela della *privacy* dei terzi ha progressivamente assunto un rilievo maggiore, divenendo oggetto di numerosi interventi normativi, sino a quello introdotto con la ricordata legge 9 agosto 2024, n. 114, con cui il legislatore ha cercato di ottenere una più pregnante tutela della riservatezza dei contenuti avulsi dal *thema probandum*¹².

Il nuovo art. 114, c. 2-*bis*, c.p.p., stabilisce il divieto assoluto di pubblicazione, anche parziale, delle intercettazioni non riprodotte dal giudice nella motivazione di provvedimenti o non utilizzate nel corso del dibattimento. L'art. 116, c. 1, c.p.p., poi, vieta il rilascio di copia delle intercettazioni di cui è vietata la pubblicazione, qualora la richiesta provenga da un soggetto diverso dalle parti o dai loro difensori. L'art. 268, c. 2-*bis*, c.p.p., infine, pone un vero e proprio "obbligo di vigilanza"¹³ in capo al pubblico ministero, tenuto a verificare che nei verbali non siano riportate espressioni lesive della reputazione, riferimenti a fatti o circostanze afferenti la vita privata degli interlocutori, nonché elementi idonei a identificare soggetti diversi dalle parti.

In precedenza, l'art. 114 c.p.p. al c. 2 *bis* consentiva la diffusione delle comunicazioni purché ritenute utilizzabili in seguito al filtro operato nella udienza stralcio. L'obiettivo, sempre il medesimo, era quello di tamponare la divulgazione di intercettazioni "aliene". Il problema, tuttavia, già emerso con la l. 7/2020, risiedeva nel fatto che all'introduzione di questo limite

¹² Le insistenti istanze di riforma volte a tutelare la riservatezza dei terzi sono, come già anticipato, alimentate anche dalla giurisprudenza della Corte EDU. Particolarmente rilevante, in tal senso, è la recentissima pronuncia CEDU, sez. I, 23 maggio 2024, n. 2507/19, in *Ius penale* (Giuffrè), con nota di D. Foti, *La disciplina italiana delle intercettazioni viola l'art. 8 CEDU*. Come segnalato dall'A., «in relazione alla disciplina delle intercettazioni, i Giudici di Strasburgo ravvisano all'unanimità la violazione dell'art. 8 CEDU poiché la legge italiana non offre a chi sia intercettato senza rivestire lo status di parte nel procedimento penale, garanzie adeguate ed effettive contro eventuali abusi. A tali soggetti, infatti, è precluso l'accesso ad un'autorità giudiziaria per promuovere il controllo sulla legittimità, proporzionalità e necessità dell'atto e, in caso di lesione del diritto al rispetto della vita privata e familiare, non può aver luogo alcuna forma di riparazione».

¹³ W. Nocerino, *Le recenti riforme in materia di intercettazioni: alcune criticità e qualche prospettiva*, cit.; C. Parodi, *La riforma Nordio e le intercettazioni: una nuova prospettiva di valutazione*, in *Ius Penale* (Giuffrè), 16 giugno 2023, sottolinea che «il perno del nuovo sistema deve essere individuato nel menzionato obbligo per il P.M. di dare indicazioni e vigilare affinché nei verbali non siano riportate espressioni lesive della reputazione delle persone o quelle che riguardano dati personali definiti sensibili dalla legge, salvo che si tratti di intercettazioni rilevanti ai fini delle indagini». L'A. si sofferma poi approfonditamente su tale obbligo di vigilanza, osservando che esso «pone a carico del P.M. una non semplice attività di vigilanza, la cui concreta attuazione potrà rivelarsi particolarmente complessa, specie nei procedimenti di maggiori dimensioni». Cfr., sul tema, anche A. Apollonio, *D.d.l. Nordio in materia di intercettazioni: l'ennesima ombra gettata sull'operato del pubblico ministero (e l'ennesimo passo verso la separazione delle carriere)*, cit., spec. par. 2, 3.

non corrispondesse un adeguato apparato sanzionatorio in caso di violazione. La pubblicazione indebita, infatti, costituiva sì reato ai sensi dell'art. 684 c.p. ma la relativa sanzione, obblazionabile con il pagamento di soli 125 euro, risultava di fatto inefficace.

Cosa cambia con la nuova riforma? La recente modifica del c. 2-*bis* dell'art. 114 c.p.p. consente la pubblicazione solo ove le stesse siano state utilizzate o nella motivazione di un provvedimento o nel corso del dibattimento¹⁴. Da oggi, dunque, si amplia in effetti il novero del non pubblicabile¹⁵.

A questo punto sarebbe lecito chiedersi il perché della “novella” e, in particolare, del requisito della concreta utilizzazione da parte del giudice. In questo modo, infatti il regime di diffusione dipende dallo stile redazionale dei provvedimenti, che non necessariamente riflette un rigoroso criterio di pertinenza o di rilevanza. Per formulare solo l'esempio più banale, ben può essere rilevante, per mostrare l'insussistenza del *fumus*, una conversazione ignorata dal giudice che tuttavia scagiona l'indagato.

L'impressione è che l'obiettivo fosse soltanto quello di ampliare le zone d'ombra, senza che la tenda del segreto fosse mossa da una mano razionale e ben orientata rispetto alla architettura generale dell'ordinamento.

Tra l'altro, non dimentichiamo che il sistema prevede il perdurare di quelle che si potrebbero a buon diritto chiamare “sanzioni pagliaccio”, ovvero il già menzionato art. 684 c.p., poiché, ancora una volta, il legislatore ha ommesso di intervenire su un aspetto cruciale del sistema, ovvero l'apparato sanzionatorio, questa volta, e per una volta, irragionevolmente mite¹⁶.

Basterebbe solo questo elemento per svelare l'inganno della riforma: i

¹⁴ Critico della scelta operata dal legislatore è, tra gli altri, A. Apollonio, *D.d.l. Nordio in materia di intercettazioni: l'ennesima ombra gettata sull'operato del pubblico ministero (e l'ennesimo passo verso la separazione delle carriere)*, cit., par. 5, 7, il quale osserva: «L'interpolazione normativa operata dal d.d.l. Nordio, questa più delle altre esaminate, conduce ad una concezione invero squalificante della figura del pubblico ministero, perché viene sancito il principio per cui, in materia di intercettazioni, vige il divieto assoluto di pubblicazione di qualsiasi atto della polizia giudiziaria (che opera sotto il controllo del pubblico ministero) e dello stesso pubblico ministero (quale ad esempio una richiesta di misura cautelare), se non viene ripreso dal giudice – e fatto proprio – nel suo provvedimento. Confrontando il dato precedente con l'attuale, può dirsi – volendo rimarcare gli effetti più radicali – che è solo l'atto di “appropriazione” del dato dell'intercettazione da parte del giudice a rendere legittima l'acquisizione dello stesso; e che il controllo effettuato *medio tempore* dal pubblico ministero conserva un valore residuale, meramente infra-procedimentale, perché lo legittima soltanto a presentare quel dato al giudice ai fini delle sue richieste».

¹⁵ Di nuovo, critico verso la scelta del legislatore è A. Apollonio, *D.d.l. Nordio in materia di intercettazioni: l'ennesima ombra gettata sull'operato del pubblico ministero (e l'ennesimo passo verso la separazione delle carriere)*, cit., par. 5, 7, il quale osserva: «È interessante notare che nei lavori preparatori del d.d.l. in commento si parla, con riferimento alla modifica dell'art. 114, di “ampliamento” del divieto di pubblicazione del contenuto delle intercettazioni. Così non è: non si tratta di una modifica quantitativa (un mero “ampliamento”), ma qualitativa: è sovvertita la natura del divieto, che adesso copre ogni atto del pubblico ministero, passando da una distinzione tra fasi del procedimento ad una distinzione tra i soggetti del procedimento: non più il pubblico ministero, ma solo il giudice valuta la rilevanza delle conversazioni captate ai fini di una eventuale e successiva ostensione».

¹⁶ Anche la Giunta dell'Unione delle Camere Penali, nella comunicazione dal titolo La Giunta UCPI sul ddl Nordio, del 15 giugno 2023, ha definito «del tutto deludente (...) la riforma in tema di intercettazioni, limitata alla questione della loro pubblicazione, senza peraltro intervenire sulla tematica principale: la sanzione di quei divieti, che rimane irrisoria e dunque di fatto inesistente».

grandi principi, gli interessi personalissimi, i diritti inviolabili garantiti dalla nostra Costituzione sembrano tutelati con una spada di cartone.

3. L'inedito divieto di pubblicazione delle ordinanze di custodia cautelare

Il d.lgs. 198/2024 è intervenuto, tra l'altro, al fine dichiarato di adeguare l'ordinamento interno alla direttiva europea 2016/343/UE. Il legislatore domestico ha così delegato il Governo a introdurre un divieto di pubblicazione dell'ordinanza di custodia cautelare sino alla conclusione delle indagini preliminari o al termine dell'udienza preliminare¹⁷. A dire il vero, il legislatore europeo non aveva espressamente toccato il tema, ma solo imposto agli Stati membri di assicurare a imputati e indagati il rispetto effettivo del principio di presunzione d'innocenza¹⁸.

Comunque sia, nel codice è stato introdotto il c. 6-ter all'art. 114 c.p.p., ai sensi del quale: «Fermo quanto disposto dal c. 7, è vietata la pubblicazione delle ordinanze che applicano misure cautelari personali fino a che non siano concluse le indagini preliminari ovvero fino al termine dell'udienza preliminare».

¹⁷ L. 15/2024, *Delega al Governo per il recepimento delle direttive europee e l'attuazione di altri atti dell'Unione europea - Legge di delegazione europea 2022-2023*, in *Gazz. Uff.*, 24 febbraio 2024, n. 46. Per un commento alla Legge si veda C. Gabrielli, *È la stampa, bellezza. Note critiche al cd. "emendamento Costa"*, in *questionegiustizia.it*, 29 febbraio 2024.

¹⁸ Dello stesso avviso è G. Giostra, *Vietare la pubblicazione dell'ordinanza di custodia cautelare: una proposta funzionalmente e sistematicamente in difficoltà di senso*, cit., 2, il quale ravvisa persino un contrasto tra la scelta operata dal legislatore nazionale e la previsione europea. Secondo l'A., infatti, «nella relazione che accompagna lo schema di decreto legislativo in oggetto si chiarisce che la modifica dell'art. 114 c.p.p. viene proposta, nel rispetto degli artt. 21, 24 e 27 della Costituzione, "in coerenza con quanto disposto dagli articoli 3 e 4 della Direttiva UE 2016/343". Ma per la verità nel decreto legislativo in esame non v'è proprio traccia dell'asserita coerenza. Anzi. L'art. 3 impegna gli Stati ad assicurare che agli indagati e agli imputati "sia riconosciuta la presunzione di innocenza fino a quando non ne sia stata legalmente provata la colpevolezza" (da intendere "provata con sentenza irrevocabile": cfr. art. 2): non solo non si può dire che il divieto *de quo* sia attuativo di una tale prescrizione, ma risulta con essa in aperto contrasto, là dove prevede che l'ordinanza si possa poi pubblicare a indagini o a udienza preliminare concluse. Se il divieto proposto fosse davvero funzionale al rispetto della presunzione di innocenza dovrebbe perdurare sino all'accertamento definitivo della colpevolezza». Inoltre, «pur scrutando in ogni interstizio [del successivo art. 4] non si riesce a trovare un posto dove alloggiare un divieto come quello che ad essa si vorrebbe ricondurre». Sul punto v. anche L. Filippi, *Ma quale presunzione di innocenza?*, in *Penale. Diritto e Procedura*, 13 gennaio 2025; Id., *Le intercettazioni dopo la legge Nordio*, 10 ottobre 2024, in *Ius Penale* (Giuffrè) ove si afferma che «la legge di delegazione quindi è andata oltre ciò che richiedeva la direttiva U.E.». Nello stesso senso A. Nappi, *Contro le virgolette: il garantismo delle perifrasi*, in *questionegiustizia.it*, 19 marzo 2024, che osserva come «la presunzione di innocenza, evocata dal richiamo all'art. 27 Cost. e alla direttiva (UE) 2016/343, peraltro già attuata con l'art. 115-bis c.p.p., non ha nulla a che vedere con il programmato "divieto di pubblicazione integrale o per estratto del testo dell'ordinanza di custodia cautelare"». Infine, anche A. Zampini, *Divieto di pubblicazione delle ordinanze che applicano misure cautelari personali: una scelta politica censurabile e costituzionalmente repressibile*, cit., sostiene che «non pare che le disposizioni europee invocate dal legislatore interessino il fenomeno della pubblicazione dei provvedimenti cautelari».

Questa scelta è stata definita «una inversione di rotta radicale»¹⁹ del legislatore che, tuttavia, in tema, la “buona rotta” sembrava averla perduta già da anni. Con il d.lgs. 216/2017, infatti, si era curiosamente sottolineato ciò su cui nessuno aveva mai sollevato dubbi, ovvero la pubblicabilità delle ordinanze *ex art.* 292 c.p.p. Tali provvedimenti non sono atti di indagine circostanza che li ha sempre esclusi dal novero di quelli coperti da segreto. Sicché, l'inciso introdotto in fine al c. 2, secondo il quale «è sempre vietata la pubblicazione degli atti [...] fatta eccezione per l'ordinanza di cui all'art. 292 c.p.p.», si riduceva a un mero *nonsense* normativo: fondato su un erroneo presupposto.

La Riforma Nordio, dunque, elimina l'inciso, ma ciò ovviamente non basta a rendere non pubblicabili le ordinanze. Dunque, con lo stesso atto normativo viene sancito espressamente il divieto di pubblicazione di tali provvedimenti, con una formulazione dai profili tanto poco ragionevoli da sollevare dubbi di incostituzionalità.

Come già evidenziato, il legislatore riconduce la recente modifica alla necessità di attuare la direttiva europea sulla presunzione di innocenza. Non si coglie, tuttavia, in che modo la pubblicazione di un'ordinanza cautelare, anche integrale, purché non travisata, possa ledere il principio in esame.

Certo, l'indagato non deve essere trattato o anche solo indicato – dentro o fuori l'aula – come colpevole prima di una sentenza di condanna passata in giudicato. In quest'ottica, la necessità che il giudice rispetti determinati criteri redazionali, chiarendo la natura meramente indiziaria degli elementi riportati e salvaguardando così il principio di presunzione di innocenza, è già richiamata dall'art. 115-*bis* c.p.p.

D'altra parte, invece, l'applicazione di una misura cautelare richiede la sussistenza di gravi indizi di colpevolezza, che devono essere puntualmente indicati a tutela e a garanzia dell'interessato. Soltanto se essi esistono, infatti, può essere giustificata la lesione della libertà personale. E in una società democratica è preciso interesse di tutti controllare che un potere così incisivo sia esercitato correttamente, circostanza che può avvenire soltanto se i provvedimenti con cui esso si esprime sono divulgabili.

Il fatto che lo siano nel contenuto e non nel virgolettato non risolve la questione, si limita a riproporre un meccanismo che ha mal funzionato in passato. Per di più, quando la libertà di una persona è compressa nel modo più grave possibile, prima di una condanna irrevocabile, è opportuno che sia verificabile nel dettaglio, “alla lettera”, se il giudice ha fatto buon governo di presupposti ed esigenze, in altri termini se questa crudele eccezione fosse davvero necessaria, ovvero corrispondesse ai requisiti previsti dall'ordinamento²⁰.

¹⁹ A. Zampini, *Divieto di pubblicazione delle ordinanze che applicano misure cautelari personali: una scelta politica censurabile e costituzionalmente repressibile*, cit.

²⁰ In proposito, già G. Illuminati, *Divieto di pubblicazione e formazione del convincimento giudiziale*, in Aa. Vv., *Processo penale e informazione. Proposte di riforma e materiali di studio*, Macerata, 2001, 52, osservava che «oltre che difficilmente attuabile sul piano pratico, da simile distinzione scaturiscono non di rado distorsioni nocive sia sul piano del diritto all'informazione sul processo che su quello di tutela di imparzialità del giudice dibattimentale. In altre parole, o il divieto di pubblicazione è totale o rischia di apparire non solo ipocrita e di conseguenza non attuato, ma addirittura controproducente per quel valore cui intenderebbe apprestare tutela».

Il fatto che si possa comunque pubblicare il contenuto dell'ordinanza mostra inoltre come l'impegno alla tutela del principio di innocenza sia solo apparente e generi un effetto paradossale: il testo dell'ordinanza redatto dal giudice nel rispetto di determinati criteri redazionali che non violino la presunzione di innocenza non potrà essere pubblicato, mentre il suo contenuto, sintetizzato da giornalisti che rischiano una cronaca accusatoria attraverso interpretazioni, magari anche distorte in buona fede, del testo, invece sì²¹.

Infine, rimane uno dei più squillanti campanelli d'allarme di quello che si annuncia come l'ennesimo fallimento dell'ultima riforma sul tema: la sanzione irrisoria (ai sensi dell'immarcescibile art. 684 c.p.) all'inosservanza del divieto.

Giungendo alle conclusioni, in tema di tutela della riservatezza, una pubblicazione non controllata del contenuto delle intercettazioni è lesiva della *privacy* dei soggetti coinvolti – che siano parti o no – nel procedimento. Tuttavia, il punto focale della questione risiede nel trovare un corretto bilanciamento tra diritto all'informazione e tutela delle libertà fondamentali. L'art. 137, c. 3, d.lgs. 196/2003 ne esprime uno condivisibile quando consente la diffusione, per finalità di informazione, di qualunque dato, purché corretto ed essenziale per comprendere una notizia di interesse pubblico. Se tale regola fosse stata applicata con rigore e con una certa solerzia, il che non è stato, forse avrebbe costituito un argine al problema dotato di una certa solidità.

Per quanto riguarda la presunzione di innocenza, la questione, sollecitata dall'Unione europea, è stata recepita dal nostro legislatore in modo, per certi versi, ambiguo.

Il problema del mancato rispetto della presunzione d'innocenza è reale, ed è acuito da una stampa spesso non all'altezza del compito, per quanto certamente arduo sia, di essere espressione di una libertà, mediatrice culturale, traduttrice della sempre maggior complessità del mondo odierno e, infine, controllore del potere.

Per essere all'altezza di tali e tanti ruoli il giornalismo dovrebbe dotarsi di strumenti culturali e professionali non comuni.

Il punto della questione, allora, non dovrebbe essere estendere il segreto in maniera pressoché indiscriminata, ma garantire un'informazione quanto più corretta ed accurata possibile. Anche perché, e questo non dovrebbe essere mai dimenticato, una volta esteso il segreto, dietro le porte chiuse

²¹ Cfr. A. Nappi, *Contro le virgolette: il garantismo delle perifrasi*, in *questionegiustizia.it*, 19 marzo 2024. Come evidenziato anche da G. Giostra, *Vietare la pubblicazione dell'ordinanza di custodia cautelare: una proposta funzionalmente e sistematicamente in difficoltà di senso*, cit., 4, è «davvero singolare che il legislatore, dopo aver opportunamente dettato regole restrittive per la redazione dell'ordinanza cautelare, in modo che non esondi senza necessità in considerazioni relative alla colpevolezza dell'indagato, ne vieti la pubblicazione lasciando però il giornalista libero di rappresentarne il contenuto senza vincolo alcuno». Secondo L. Filippi, *Ma quale presunzione di innocenza?*, cit., 4, «paradossalmente il cronista giudiziario, non vincolato dal testo, acquista maggiore libertà nell'esercizio della cronaca giudiziaria»; Id., *Le intercettazioni dopo la legge Nordio*, cit., 3, rileva che «si registra così un *revirement* solo apparente: l'ordinanza cautelare e l'informazione di garanzia saranno pubblicabili liberamente interpretate dal cronista, riassunte o commentate: il che è molto pericoloso perché il cronista scrive sugli atti di indagine del solo P.M. e inoltre può esagerare e addirittura travisare il contenuto dell'ordinanza».

dei tribunali tutto può accadere. Il potere, una volta celatosi allo sguardo esterno, ha sempre la tendenza a degenerare²².

Ciò non significa che la diffusione di informazioni sul processo non possa essere spiacevole, se non lesiva, per chi ne è coinvolto. Ma, a livello di sistema, sembrano più gravi i rischi di una giustizia segreta rispetto ai pericoli di una stampa il più libera (e plurale) possibile, anche tenuto conto di occasionali stravolgimenti o malintesi nella interpretazione dei provvedimenti e persino nel caso del resoconto di provvedimenti poi annullati²³. D'altronde, a volte, una corretta informazione può rivelarsi un volto amico per l'indagato nello svelare eventuali storture dell'indagine.

Al di là delle soluzioni discutibili da un punto di vista di tecnica legislativa, quello che oggi appare malinconicamente chiaro è che l'attuale governo in veste di legislatore sembra avere cambiato idea, per così dire, sul ruolo della stampa nel regolare il bilanciamento fra i vari interessi in gioco quando si tratta di disegnare le regole di dettaglio sulla pubblicabilità degli atti del processo penale.

I giornalisti, da rappresentanti di un contrappeso indispensabile per controllare il potere, sono visti come mero veicolo di pettegolezzo. Col che ci si spiega la ragione per cui le disposizioni non mirano a consentire la più ampia pubblicazione, purché ciò non configga con il corretto svolgimento del processo e con un razionale bilanciamento con i diritti della persona, ma tendono a vietare il più possibile, senza nemmeno preoccuparsi troppo del fatto che la latitudine dei divieti abbia una razionalità in sintonia con l'architettura costituzionale.

Nessuno dubita che l'informazione giudiziaria nel nostro Paese goda di cattiva salute. Tuttavia, continuiamo a restare affezionati all'idea che limitarsi a espandere il segreto da un lato non risolva davvero il problema, dall'altro abbia controindicazioni notevoli e persino pericolose per la sopravvivenza del sistema come lo conosciamo.

Insomma, uccidere il paziente, non pare un buon modo per debellare la malattia.

²² Si coglie, in questa prospettiva, l'esigenza di rendere effettivo il ruolo della stampa quale «cane da guardia» della democrazia. E perché tale funzione risulti effettivamente svolgibile, essa deve poter operare con quella libertà di movimento che contraddistingue – per riprendere l'immagine – «il buon cane da guardia [che] gira libero attorno a casa, orecchie tese e naso al vento. E abbaia, anche più forte del necessario e qualche volta deve mordere. Così la stampa» (V. Zagrebelsky, *Se il cane da guardia non morde*, in *La Stampa*, 30 settembre 2011).

²³ Come rilevato da G. Giostra, *La giustizia penale nello specchio deformante della cronaca giudiziaria*, cit., 26, «il valore della pubblicità della giustizia penale, quindi, non va misurato soltanto sugli effetti che essa in concreto propizia, ma sulla gravissima involuzione civile e democratica che la sua assenza comporterebbe. Sarebbe quindi costituzionalmente, politicamente e culturalmente inammissibile oscurare la cronaca giudiziaria».

Abstract

Il contributo esamina le recenti evoluzioni normative in tema di pubblicazione degli atti del processo penale, con particolare attenzione alle intercettazioni e alle misure cautelari. L'analisi mette in luce le tensioni tra tutela della riservatezza, presunzione di innocenza e diritto all'informazione, evidenziando le criticità di un ampliamento eccessivo delle aree di segretezza e i suoi effetti sul controllo pubblico dell'attività giudiziaria.

The article addresses recent developments in the rules governing the publication of criminal procedure documents, with a focus on wiretapping materials and pre-trial measures. It explores the tensions between privacy protection, the presumption of innocence, and freedom of information, highlighting the risks posed by an excessive expansion of secrecy for public scrutiny of judicial activity.

Keywords

cronaca giudiziaria - giornalismo - diritto di informazione - processo penale

La Dichiarazione per la sovranità digitale europea e l'utilità (trascurata) del ricorso alla geopolitica*

Paolo Passaglia

Sommario

1. Introduzione. – 2. L'adozione della Dichiarazione per la sovranità digitale europea. – 3. La genericità della Dichiarazione – 4. ... e l'opportunità di un'analisi geopolitica dei fenomeni. – 5. Un appunto finale.

1. Introduzione

Nel contesto attuale di instabilità e incertezza che affligge (anche) le società occidentali, alcune nozioni ricorrono con sempre maggiore frequenza, ora come scopo cui tendere ora come punto fermo da stabilire: tra queste, una posizione non marginale è da riconoscere alla c.d. «sovranità digitale»¹. Non è ben chiaro da dove l'espressione derivi², né può dirsi che cosa esattamente riassume, anche perché le definizioni che se propongono sono apprezzabilmente disparate. Quale che sia il significato che si adotti, è peraltro indubbio che l'utilizzo del sostantivo sia quanto meno ardito,

* L'articolo è stato sottoposto, in conformità al regolamento della Rivista, a referaggio anonimo.

¹ La diffusione dell'espressione trova riscontri evidenti anche in chiave bibliografica. Limitandosi ad alcuni recenti contributi, tra i più noti si può fare riferimento a B. Farrand – H. Carrapico, *Digital sovereignty and taking back control: from regulatory capitalism to regulatory mercantilism in EU cybersecurity*, in *European Security*, 3, 2022, 435 ss.; H. Roberts – E. Hine – L. Floridi, *Digital Sovereignty, Digital Expansionism, and the Prospects for Global AI Governance*, in M. Timoteo – B. Verri – R. Nanni (eds.), *Quo Vadis, Sovereignty?*, Cham, 2023, 51 ss.; S. Fratini – E. Hine – C. Novelli – H. Roberts – L. Floridi, *Digital Sovereignty: A Descriptive Analysis and a Critical Evaluation of Existing Models*, in *Digital Society*, 3, 2024, 59 ss.; P. Bellanger – A.-T. Norodom, *La souveraineté numérique*, in M. Caron – R. Maurel (sous la direction de), *Penser la transition numérique. Vers un monde digital durable*, Ivry-sur-Seine, 2023, 41 ss.; B. Brunessen – G. Le Floc (sous la direction de), *La souveraineté numérique*, Bruxelles, 2024; in lingua italiana, v., ora, R. Baldoni, *Sovranità digitale. Cos'è e quali sono le principali minacce al cyberspazio nazionale*, Bologna, 2025.

² La paternità dell'espressione viene generalmente attribuita, specie in ambienti francofoni, a Pierre Bellanger, fondatore del gruppo radiofonico Skyrock, che avrebbe fatto uso di questo concetto già nei primi Anni Duemila, per poi esplicitarlo e definirlo nel decennio scorso (v., in part., P. Bellanger, *De la souveraineté en général et de la souveraineté numérique en particulier*, in *Les Echos*, 30/08/2011; Id., *La souveraineté numérique*, Paris, 2014).

mentre l'aggettivo può risultare spiazzante, non fosse altro perché appare sempre più difficile tracciare la distinzione tra ciò che è «digitale» e ciò che è esterno a questa dimensione: le intersezioni e le sovrapposizioni tra mondo fisico e mondo virtuale, tra «digitale» e «analogico», sono divenute un aspetto consustanziale della vita collettiva e di quella individuale, come hanno cercato di dimostrare³, anche sul piano giuridico, i molti che hanno teorizzato un diritto di accesso alle nuove tecnologie, ma anche coloro che hanno teorizzato un diritto a non utilizzarle⁴.

Trascurando queste problematiche – che pure sarebbero di estrema importanza – per focalizzarsi sul termine «sovranità», viene quasi necessitato il qualificare l'utilizzo in termini ottativi, ben più che realisticamente descrittivi: se si prendono in considerazione i più classici attributi della sovranità, appare evidente che siano tutti difficili da identificare all'interno dell'ecosistema digitale.

Si pensi alla natura originaria dell'ente sovrano, che nel contesto delle nuove tecnologie lascia il posto a un'affermazione del potere in buona misura derivata da ciò che le nuove tecnologie stesse permettono. Non è infatti il sovrano ad auto-fondare il proprio potere, ma sono piuttosto i progressi tecnologici a tracciare (a seconda dei casi, in senso espansivo o in senso restrittivo) il quadro entro il quale il «sovrano» può muoversi.

Analoghe riserve possono essere esposte in merito al carattere essenziale della sovranità interna, e cioè la supremazia rispetto a qualunque altro soggetto: in questo caso, la sovranità riferita a un qualunque ente deve fare i conti con la natura perlomeno frastagliata dell'ecosistema digitale, in cui – tra l'altro – un ruolo determinante è giocato da soggetti privati, nei confronti dei quali si è chiaramente in difficoltà a dispiegare una vera supremazia (sia che si manifesti in termini di potere regolativo o sia che si esprima nel monopolio della forza).

Finanche sotto il profilo esterno della sovranità, quello dell'indipendenza dagli altri enti sovrani, emergono difficoltà non trascurabili, dal momento che qualunque soggetto o ente che ambisca a definirsi indipendente rischia di scontrarsi con le evidenze di segno opposto, quelle che connotano una interdipendenza difficile da districare (nel migliore dei casi) o (peggio) quelle che conducono all'inevitabile constatazione dell'impossibilità pratica di superare condizionamenti provenienti dall'esterno.

Poste queste premesse, il richiamo al concetto di «sovranità digitale», più che andare a identificare determinate forme di controllo e di gestione dell'ecosistema digitale, sembra che svolga una funzione quasi «rassicurante», nel senso di disegnare un consolidamento di poteri che si percepiscono come minacciati oppure un orizzonte di impegno finalizzato a riaffermare una presenza non puramente passiva in un contesto più ampio⁵.

³ Sia consentito rinviare, per una sintesi e per i dovuti riferimenti bibliografici, a P. Passaglia, *La problematica definizione dell'accesso a Internet e le sue ricadute su esclusioni sociali e potenziali discriminazioni*, in questa Rivista, 3, 2021, 125 ss.

⁴ Sul tema, v., di recente, D. Kloza – E. Kuzelewska – E. Lievens – V. Verdoodt (edited by), *The Right Not to Use the Internet. Concept, Contexts, Consequences*, London, 2025.

⁵ Emblematica, a tal riguardo, è la definizione, per così dire «fondativa» (v. *supra*, nota 2), che Bellanger ha proposto della sovranità digitale, dove il punto centrale è proprio quello di poter «avere in mano» il proprio presente e il proprio destino: *«la maîtrise de*

Non è un caso, del resto, che se il concetto astratto viene sovente evocato, non è dato individuare un ente che si definisca «sovrano digitale»: nessuno lo è, alcuni (o magari anche molti) vorrebbero esserlo, tutti sanno di non poterlo essere o di non poterlo essere davvero fino in fondo. L'eccezione è rappresentata da questi riferimenti alla sovranità che, evocando veri o presunti pericoli esterni, sono finalizzati a creare confini protetti all'interno dei quali è possibile esercitare più o meno stringenti forme di censura: qui il sovrano effettivamente esiste, ma come aggiornamento in tempi di rivoluzione digitale di forme premoderne di esercizio del potere⁶.

2. L'adozione della Dichiarazione per la sovranità digitale europea

Tenendo fermo il carattere eccezionale della fattispecie da ultima ricordata, è, probabilmente, utilizzando la prospettiva di analisi subito sopra proposta che debbono essere interpretati quei documenti, quelle affermazioni e quegli atti nei quali, a nome di enti «analogicamente sovrani», ci si investe anche della sovranità digitale.

Ultima in ordine di tempo, ma non di certo in ordine di importanza, è la *Declaration for European Digital Sovereignty*, adottata a Berlino il 18 novembre scorso dai rappresentanti degli Stati membri dell'Unione europea⁷. Si tratta di un testo che si colloca al di fuori del diritto euro-unitario, ponendo obiettivi e principi che prescindono dai confini competenziali, e che quindi non si riferiscono solo all'Unione, ma anche agli Stati. Si tratta, però, è bene precisarlo chiaramente, di un testo che pone, sì, obiettivi e principi generali, ma che non reca norme vincolanti: a tutto concedere, può dunque ambire al rango di *Soft Law*, senza che possa escludersi (anzi...) che finisca per essere uno dei non pochi esempi di mere declamazioni, dall'efficacia concreta non esattamente assicurata⁸.

Ora, nonostante un'importanza che, sul piano pratico, è senz'altro da relativizzare in maniera vistosa, la Dichiarazione in discorso merita di essere analizzata, non fosse altro perché si pone come il precipitato istituzionale

notre présent et de notre destin tels qu'ils se manifestent et s'orientent par l'usage des technologies et des réseaux informatiques» (P. Bellanger, *De la souveraineté en général et de la souveraineté numérique en particulier*, cit.).

⁶ L'esempio forse più eloquente viene dalla Russia, dove la legge federale n. 90-FZ del 1° maggio 2019, recante «Modifiche alla legge federale “sulle comunicazioni” e alla legge federale “sull'informazione, sulle tecnologie dell'informazione e sulla protezione dell'informazione”», che ha creato un sistema centralizzato di controllo del traffico Internet nazionale, viene informalmente definita come la «Legge sull'Internet sovrano».

⁷ Il testo, in inglese, della Dichiarazione è consultabile anche sul sito del Dipartimento per la trasformazione digitale del Governo italiano, in allegato all'articolo dal titolo *Dichiarazione Sovranità Digitale Europea: Italia sottoscrive il documento con gli altri Stati membri*, 19 novembre 2025.

⁸ Sulla distinzione tra *Hard Law* e *Soft Law*, v. E. Mostacci, *La soft law nel sistema delle fonti: uno studio comparato*, Padova, 2008; A. Somma (a cura di), *Soft law e hard law nelle società post-moderne*, Torino, 2009; G.C. Shaffer – M.A. Pollack, *Hard Vs. Soft Law: Alternatives, Complements, and Antagonists in International Governance*, in *Minnesota Law Review*, 2010, 712 ss.

di una serie di iniziative che si sono sviluppate negli ultimi anni, soprattutto dopo il periodo dell'emergenza pandemica⁹.

Un primo aspetto da sottolineare è la definizione di sovranità digitale che viene fornita, vale a dire: «la capacità dell'UE e dei suoi Stati membri di agire autonomamente e di scegliere liberamente le loro proprie soluzioni, raccogliendo al contempo i benefici della collaborazione con *partners* globali, quando possibile»¹⁰. Se le parole hanno un senso, una tale definizione ha ben poco a che vedere con il concetto di «sovranità», ponendo in evidenza, piuttosto, quello di «autonomia», collegata a una libertà (e non a un potere) di scelta.

Una dichiarazione, che si vuole quanto più solenne possibile, la quale reca un *incipit* di questo tipo è già di per sé piuttosto indicativa dello scollamento tra l'«essere» e il «voler essere». Sin da subito, del resto, si evocano soggetti con i quali collaborare, «se possibile»: soggetti che sono globali, e che quindi operano a un livello più ampio dell'Unione. La constatazione di questa diversità di livelli è un chiaro avvertimento inerente a una posizione di potenziale minorità dell'Europa, tanto è vero che si ammette che la collaborazione non sia scontata, il che significa, come minimo, che l'Unione non è in grado di imporla.... D'altra parte, che non si tratti di «sovranità» è scritto in maniera ancora più nitida nel periodo immediatamente successivo a quello citato, nel quale il potere di scelta delle proprie soluzioni si dice implicare «la definizione e l'applicazione di una regolamentazione giuridica propria, conforme ai principi accettati a livello internazionale»¹¹: in definitiva, la Dichiarazione non è un'affermazione di sovranità, bensì una rivendicazione di autonomia in un contesto internazionale in cui operano attori globali.

E la sostanza non cambia neppure quando, subito dopo, nel rifiutare isolazionismo e protezionismo, si sostituisce al concetto di «autonomia» quello di «indipendenza»; perché non basta dire che l'Europa deve poter «agire indipendentemente e in maniera autodeterminata», se poi, nell'enucleare i canoni su cui basare la propria azione, si parte dal «diritto internazionale», e solo dopo si citano «i propri diritti, valori e interessi di sicurezza»¹². Un reale recupero di «sovranità» non può certo venire da questo; più efficace, semmai, è la precisazione dell'indirizzo da dare alla cooperazione internazionale, che si vuole sviluppare con i *partners* «che condividono i valori e i principi europei»¹³: almeno, non si tratta di una rassegnata ricerca della

⁹ Per una efficace ricostruzione delle politiche poste in essere dall'Unione europea in vista dell'affermazione di una sovranità digitale, v. l. von Ditzfurth, *The European Union's Pursuit of Digital Sovereignty Through Legislation*, in *Journal of Intellectual Property, Information Technology and Electronic Commerce Law*, 2, 2025, 286 ss.

¹⁰ «[...] the EU and its Member States' ability to act autonomously and to freely choose their own solutions, while reaping the benefits of collaboration with global partners, when possible»: Declaration for European Digital Sovereignty, 1° cpv.

¹¹ «[The digital sovereignty] involves setting and enforcing our own legal framework in accordance with internationally accepted principles»: *ibid.*

¹² «Digital sovereignty does not mean isolation or protectionism; it means ensuring that Europe can act independently and in a self-determined manner based on international law, its own laws, values, and security interests»: *ibid.*

¹³ «[...] thriving to international cooperation with its partners that share European values and

collaborazione con chiunque.

In effetti, in modo molto più concreto, la definizione successivamente fornita di sovranità digitale diventa la capacità dell'Europa, degli Stati e dei soggetti in essi operanti «ad agire in modo indipendente nel mondo digitale, rendendo possibili decisioni autonome circa l'utilizzo, la *governance* e lo sviluppo di sistemi digitali, senza un eccessivo affidamento ad attori esterni al fine di proteggere le nostre democrazie europee e i nostri valori europei»¹⁴. Non ci si può esimere dal registrare un concetto di sovranità piuttosto anomalo, che mira alla protezione dei cardini dei sistemi senza affidarsi in maniera «eccessiva» ad attori esterni. Come dire che ciò che si rifugge non è, di per sé, l'affidarsi a interventi esterni per garantire i tratti fondamentali del proprio modo di essere; è l'eccesso di aiuto richiesto a essere stigmatizzato. Se di sovranità si vuol proprio parlare, la nozione che viene per prima alla mente è allora quella degli «Stati a sovranità limitata» di brezneviana memoria...¹⁵.

Con queste premesse, i principi mediante i quali si mira a dotarsi di un «quadro di riferimento per rafforzare la capacità dell'Europa di agire nella sfera digitale, garantendo al contempo coerenza, equilibrio e apertura»¹⁶, pur essendo tutti, in sé e per sé, pianamente condivisibili, non possono non vedere la propria portata considerevolmente ridimensionata alla luce dell'obiettivo minimalista che sono chiamati a perseguire¹⁷. Del resto, non può non sottolinearsi che la Dichiarazione, già dal suo titolo, indica un fine

principles»: ibid.

¹⁴ «[The digital sovereignty] encompasses the ability of individuals, businesses and institutions in Europe to act independently in the digital world, allowing for autonomous decisions about the use, governance, and development of digital systems without undue reliance on external actors in order to protect our European democracies and our European values»: *id.*, 2° cpv.

¹⁵ Sulla c.d. dottrina Breznev, formulata nel 1968 in chiave di subordinazione dei Paesi dell'Europa centrale e orientale all'Unione sovietica, v. S.G. Glazer, *The Brezhnev Doctrine*, in *International Lawyer*, 1971, 169 ss.

¹⁶ «[These principles] provide a framework for strengthening Europe's capacity to act in the digital sphere, while ensuring consistency, balance, and openness in our approach»: *Declaration for European Digital Sovereignty*, 4° cpv.

¹⁷ A dire il vero, più che «principi», vengono enucleati obiettivi e linee di condotta, nel quadro di alcuni generalissimi principi; si elencano, infatti, le seguenti voci: semplicità e chiarezza degli interventi; creazione di un contesto che supporti gli investimenti; sovranità sui dati; visione di lungo termine e investimenti in aree strategiche; utilizzo di soluzioni *open-source*; creazione di risorse comuni a livello europeo; inventariazione delle iniziative esistenti; assunzione di un ruolo attivo nel conformare una politica internazionale che sfrutti i punti di forza dell'Europa; gestione strategica, responsabile e basata sul rischio delle dipendenze; *governance* forte; investimenti nella formazione, nella ricerca, nello sviluppo di competenze digitali e nell'alfabetizzazione digitale; rafforzamento della resilienza attraverso lo sviluppo di forti capacità e competenze in materia di cybersicurezza; garanzia di mercati equi, efficienti, basati sulla fiducia e competitivi; protezione della democrazia e rafforzamento della fiducia all'interno delle società. Da notare è la collocazione soltanto all'ultimo posto della preservazione delle istanze democratiche, mentre non figura proprio nell'elenco la protezione dei diritti fondamentali, che invece da tempo caratterizza fortemente la politica dell'Unione europea (anche) nell'ambito dell'ecosistema digitale, e che anzi di solito giustifica la sua attività regolatoria, assai più sviluppata di altri attori geopolitici. Per una critica incentrata proprio su quest'ultimo punto, v. M. Czerniawski, *EU's Digital Sovereignty and the Rights-Based Imperative: Linking Enforcement, Competences and Fundamental Rights*, in *Verfassungsblog*, 3 dicembre 2025.

(«per», *for*), che, come tale, non è detto che sia raggiungibile.

Non è tanto questo, però, il punto più delicato della Dichiarazione: in fondo, sarebbe ingeneroso criticare la modestia degli obiettivi, se non altro nella misura in cui essa discenda da iniezioni di forti dosi di sano realismo. L'aspetto su cui è forse più interessante soffermarsi attiene allo iato che si è venuto a creare tra l'approccio definitorio generale e la concretizzazione che, attraverso le condotte auspiccate, si è proposta. Per meglio dire, nel delineare i confini e le forme di manifestazione dei c.d. «principi», i redattori della Dichiarazione, che pure *in limine* si sono mostrati (anche fin troppo) consci della duplicità di scale geografiche su cui operano gli attori nazionali ed europei, da un lato, e quelli globali, dall'altro, hanno poi optato per un'analisi priva di ancoraggio al dato *lato sensu* spaziale, preferendo affermazioni generali. Affermazioni che, nella loro astrattezza, da generali sono divenute sotto molti profili generiche e, quindi, in parte non irrilevante inutili.

3. La genericità della Dichiarazione ...

Il quadro all'interno del quale ha tratto origine la Dichiarazione per la sovranità digitale è connotato, tra l'altro (ma forse soprattutto), da un duplice riconoscimento: per un verso, quello della dimensione relativa dell'Unione europea rispetto alle dinamiche globali e, per l'altro, quello delle difficoltà che l'Unione europea incontra nel fronteggiare e, *a fortiori*, nell'orientare tali dinamiche.

Alla luce di questa doppia constatazione, ci si sarebbe potuti attendere che la Dichiarazione mirasse a proporre chiavi di lettura volte al superamento almeno parziale delle indicate difficoltà, onde dar corpo a qualcosa di non troppo distante da una «sovranità», ovviamente entro i confini dell'Unione. L'elenco delle azioni proposte e dei principi da seguire, invece, si apparenta più all'espressione nel mondo digitale di talune delle caratteristiche proprie dei sistemi giuridico-istituzionali europei che non alla prospettazione dei cardini su cui impostare una reale *policy* attuativa della sovranità digitale. Ora, affinché la sovranità digitale abbia una qualche concretizzazione, nonostante gli ostacoli che ne minano in radice e irrimediabilmente il raggiungimento, è necessario che, nel momento in cui se ne progetta l'attuazione, si parta proprio da questi ostacoli, con l'idea di almeno in parte aggirarli. Ciò posto, quando si affronti la tematica della sovranità digitale, l'aspetto inevitabilmente più delicato risiede proprio nella diversità di piani tra l'Unione europea e i fenomeni che la Dichiarazione stessa definisce come «globali» e che hanno un rilievo primario nel campo delle nuove tecnologie.

Colpisce che, evidenziate, sin da subito, le difficoltà derivanti da questa discrasia di piani, la Dichiarazione, quando allude all'esistenza di «*partners globali*», assume i fenomeni che trascendono l'Unione come un dato unitario, o meglio, come un insieme di fenomeni che possono essere trattati unitariamente. Si tratta di una semplificazione che è, ad un tempo, eccessiva e fuorviante: eccessiva, perché non tiene conto dell'esistenza di differenze anche profonde tra i fenomeni variamente collegati all'ecosiste-

ma digitale; fuorviante, perché proprio l'assimilazione di fenomeni diversi impedisce di adottare un approccio poliedrico nel dar corpo alla (anelata) sovranità digitale.

Di contro, la duplicità di piani intuita *ab initio* avrebbe potuto e dovuto essere sviluppata in maniera assai più articolata e convincente; e ciò sarebbe risultato possibile se l'analisi di marca politico-istituzionale fosse stata arricchita da una visione di stampo geopolitico.

4. ... e l'opportunità di un'analisi geopolitica dei fenomeni

Quest'ultimo termine – di recente oggetto di un utilizzo che non è forse eccessivo definire inflazionistico – richiede una precisazione. In luogo della definizione che spesso di geopolitica viene fornita, e che la rende prossima all'analisi delle relazioni internazionali¹⁸, una diversa lettura può apparire più fruttuosa, nella misura in cui permette di associare la decisione *lato sensu* politica e la dimensione geografica, identificando quelle *policies* in cui il dato geografico esercita un peso non trascurabile nell'ambito del *decision-making*. Il riferimento va, essenzialmente, all'inquadramento che della disciplina ha proposto, tra la fine degli Anni Settanta e gli Anni Ottanta, Yves Lacoste¹⁹, iniziatore di una scuola ancora oggi molto seguita in Francia, coagulatasi principalmente attorno alla rivista *Hérodote*²⁰.

Ai presenti fini, l'attenzione alla geografia²¹ è particolarmente importante soprattutto perché consente di collocare su diversi ordini di grandezza i vari fenomeni che sono collegati alla sovranità digitale. Per dirla in termini più espliciti, l'ecosistema digitale non si compone soltanto di fenomeni globali, poiché al suo interno convivono fenomeni che sono effettivamente globali e fenomeni che hanno una portata geografica più circoscritta, anche molto più circoscritta, fino a essere eminentemente locali. Se l'immagine più suggestiva è quella del cyberspazio nel quale barriere e confini tradizionali evaporano in una dimensione a-spaziale²², non può trascurarsi

¹⁸ Un'analisi dell'evoluzione del termine e delle diverse angolature che la disciplina ha assunto nel corso del tempo è stata recentemente proposta anche dalla dottrina giuridica: cfr. R. Ibrido, *Geopolitica costituzionale. Una introduzione*, Bologna, 2025, spec. 19 ss.; ivi anche un'ampia bibliografia. V., volendo, pure il più sintetico P. Passaglia, *Geopolitica e diritto. Un legame trascurato*, Milano, 2026, 11 ss.

¹⁹ Sebbene non faccia impiego del termine «geopolitica», il testo di riferimento è senz'altro Y. Lacoste, *La géographie, ça sert, d'abord, à faire la guerre*, Paris, 1976.

²⁰ Attiva dal 1976, la rivista *Hérodote. Revue de Géographie et de Géopolitique*, è giunta, al termine del 2025, al suo n. 199.

²¹ Cfr. Y. Lacoste, *La géographie, la géopolitique et le raisonnement géographique*, in *Hérodote*, 130, 2008, 17 ss.

²² Tanto da suggerire a John Perry Barlow la celebre Dichiarazione di indipendenza del cyberspazio, nella quale si teorizza una dimensione che doveva restare immune agli obblighi e ai vincoli che gli Stati, quando la Dichiarazione veniva redatta (8 febbraio 1996), avevano iniziato timidamente a porre. La *Declaration of the Independence of Cyberspace* è ancora oggi consultabile sul sito della *Electronic Frontier Foundation*, sebbene il suo afflato libertario sia stato ormai ampiamente superato da una abbondante produzione normativa, nazionale e sovranazionale.

che questa dimensione, per esistere, necessita innanzi tutto di infrastrutture, che hanno invece una loro incontestabile materialità, e che sovente impongono scelte di *policy* ben radicate su territori specifici. Senza *server* e cavi non si avrebbe il cyberspazio²³, ma la collocazione degli uni e degli altri è il prodotto di opzioni nelle quali entrano in gioco fattori anche molto lontani dal mondo digitale. Allo stesso modo, lo sviluppo dell'intelligenza artificiale non sarebbe possibile senza *data centers*, la cui localizzazione, tuttavia, rappresenta una scelta geograficamente tra le più delicate, come dimostrano – tra le altre cose – l'elevato utilizzo di acqua per il raffreddamento e il livello di rumore prodotto, scarsamente compatibile con la vicinanza a centri abitati²⁴.

È chiaro che affermare una sovranità digitale deve significare, innanzi tutto, propugnare l'esercizio in piena indipendenza di poteri collegati a fenomeni rientranti integralmente all'interno del territorio in cui si esercita la sovranità in senso proprio. Non è significativo, ai presenti fini, se la competenza sia da imputare all'Unione europea o agli Stati membri: ciò che rileva è che l'ente cui il potere è attribuito sia in grado di esercitarlo.

Sul punto non conviene arrestarsi al dato puramente formale, che suggerirebbe conclusioni tanto scontate, quanto alla fin fine inutili: una vera sovranità digitale dovrebbe, infatti, essere in grado di condizionare (anche) gli attori globali in merito alle decisioni che abbiano una rilevanza tutta interna ai confini del decidente; le capacità di influenza e di condizionamento che le grandi multinazionali esercitano in molti paesi del Sud globale, anche nell'adozione di decisioni di questo tipo, dovrebbero essere un monito contro possibili espansioni di questo *decision-making* alterato verso il Vecchio Continente. E allora, una qualche presa di posizione su un aspetto tanto delicato, all'interno di una Dichiarazione per la sovranità digitale europea, non avrebbe di certo guastato, anzi sarebbe stata probabilmente più che opportuna.

L'omissione di ogni riferimento per così dire «localistico» testimonia della prospettiva unicamente «globale» nella quale è stata declinata la rivendicazione di sovranità digitale. La scelta, come detto, è opinabile in sé, per tutto quanto lascia impregiudicato, ma, a ben vedere, risulta fallace anche sotto un altro punto di vista, forse altrettanto importante. La Dichiarazione in esame fa ampio ricorso al termine «globale», assumendo che esso sia, per sua natura, descrittivo di una realtà univoca. A ben vedere, invece, i richiami sembrano semplicemente indicare l'esistenza di attori la cui sfera di attività supera quella dell'Unione europea, e che, forse, potrebbero più proficuamente definirsi con il meno impegnativo attributo della transnazionalità: come è ovvio, infatti, non tutto ciò che supera la dimensione europea è «globale». Ma ciò che più conta è che l'impiego di quest'ultimo aggettivo non consente di distinguere le diverse manifestazioni della «globalità», *recte* della «transnazionalità».

Perché le manifestazioni della transnazionalità non vadano a confluire in

²³ V., in proposito, anche per ulteriori richiami, L. Martino, *Sovranità digitale e competizione geopolitica nel contesto dei cavi sottomarini: analisi comparata degli approcci di Cina, Stati Uniti e Unione europea*, in questa Rivista, 2, 2024, 144 ss.

²⁴ Al riguardo, v. A. Aresu, *Geoeconomia dei data center*, in *Limes. Rivista italiana di geopolitica*, 11, 2024, 159 ss.

un blocco indistinto, l'approccio di tipo geopolitico diventa indispensabile, e particolarmente perspicuo è il ricorso a metodi cartografici che permettano di raffigurare i fenomeni, concretizzandone i caratteri essenziali dalla prospettiva geografica²⁵.

In effetti, un fenomeno che ha valenza transnazionale, o, al limite, anche globale, quando venga raffigurato su una mappa, può rivelare conformazioni significativamente diverse. Questo fenomeno può essere autenticamente globale, nel senso di essere diffuso in tutto il pianeta in forme almeno significative: il cyberspazio o l'uso degli *smartphones* sono due elementi che possono senz'altro annoverarsi all'interno di questa categoria. Al contrario, altri fenomeni sono presenti, magari, anche in tutte le aree del pianeta, ma la loro diffusione, lungi dall'essere uniforme, lascia emergere forti discrepanze, al punto da poter individuare aree di marcata concentrazione: una mappa del valore economico del ricorso all'*eCommerce* rende evidente che sono le aree più ricche del pianeta quelle in cui il fenomeno si concentra; una mappa della diffusione dei droni fa, invece, emergere un tasso di penetrazione particolarmente elevato nelle zone di guerra; e così via.

Queste considerazioni possono apparire, di primo acchito, del tutto estranee alla prospettiva teorica di ampio respiro che non può non essere propria di una Dichiarazione che abbia ad oggetto la sovranità digitale europea. In realtà, il loro impatto anche sulla dimensione autenticamente giuridico-istituzionale è di primaria importanza.

Se l'Unione europea intende davvero esercitare la propria sovranità digitale (o almeno provare a farlo), ad esempio adottando norme che abbiano a oggetto fenomeni transnazionali, deve, chiaramente, mettere in conto le problematiche prettamente giuridiche legate alla diversità di piani in cui ci si muove (e che dunque possono associarsi a situazioni di extra-territorialità, a esigenze di scelta del diritto applicabile, etc.); allo stesso tempo, però, è più che opportuno che non trascuri gli elementi geopolitici, perché è fondamentale da questi che può cercare di prevenire la propria forza di condizionamento.

Ne è prova il fatto che l'Unione europea, in concreto, può contare su un grande potere quando deve regolamentare fenomeni che sono transnazionali, ma che sono concentrati in poche aree del pianeta, e tra queste l'Europa ha uno spazio rilevante. I destinatari della regolamentazione sanno, infatti, di non potersi contrapporre frontalmente all'Unione europea se non a rischio di perdere una quota notevole dei propri consumatori o del proprio giro d'affari.

All'opposto, il potere di condizionamento e l'efficacia della normativa si riduce grandemente quando l'Unione europea regolamenti fenomeni transnazionali di portata autenticamente globale. In questo caso, l'esercizio delle attribuzioni inerenti alla sovranità digitale non può avvalersi

²⁵ Di nuovo fondamentale appare l'insegnamento di Lacoste, ideatore del c.d. diatopo, cioè di un metodo di mappatura consistente in «un tipo di costruzione schematica formata dalla sovrapposizione delle carte e delle rappresentazioni spaziali collocate in prospettiva in maniera da accavallarsi» (cfr. Y. Lacoste, *La géopolitique par les cartes. La longue histoire d'aujourd'hui*, Paris, 2022, 24). L'illustrazione complessa nella quale si collegano i fenomeni rintracciabili ai vari livelli rende possibile, attraverso le mappe, un'analisi dalla quale si possono più compiutamente cogliere i rapporti tra i vari fenomeni e tra questi e i soggetti cui si riferiscono.

della «rendita di posizione» riscontrata nell'ipotesi precedente. Questo può condurre a ovvie difficoltà di *enforcement*, ma la conseguenza potenzialmente più significativa consiste nella possibilità che si producano processi di «*foot voting*»²⁶. In effetti, se un certo fenomeno è globale, i destinatari delle norme dell'Unione europea possono prendere in considerazione l'idea di sfuggire alla regolamentazione introdotta subendone conseguenze relativamente limitate.

Solo per prospettare un esempio che possa essere quanto più possibile esplicativo, il caso della regolamentazione dell'intelligenza artificiale risulta particolarmente idoneo: poiché l'Europa non ha una posizione privilegiata rispetto ad altri soggetti geopolitici (non investe di più di altri nella ricerca, non ha infrastrutture più competitive, etc.), l'introduzione di una normativa rigorosa può avere l'effetto di far spostare i soggetti implicati nello sviluppo dell'intelligenza artificiale verso altre aree del pianeta, dove le loro attività sono meno limitate. Ne consegue una crisi di effettività della normativa europea, ma anche un rafforzamento di *competitors* che, in futuro, rischiano di essere in una posizione di forza tale da imporre all'Unione europea determinate politiche²⁷.

5. Un appunto finale

Il caso da ultimo prospettato suggerisce più di qualche apprensione, perché rende chiaro che la rivendicazione e il concreto esercizio di attribuzioni sovrane, qualora non vengano adeguatamente calibrati alla luce del contesto in cui si collocano, rischiano di avere pesanti effetti pregiudizievole, al punto che una affermazione di sovranità può alla lunga tradursi, addirittura, nell'origine di una posizione di debolezza.

Questo dimostra che un approccio che sia di tipo unicamente giuridico-formale alle tematiche inerenti all'ecosistema digitale, e segnatamente a quelle più «impegnative», non può mai essere adeguato. Niente, infatti, pregiudica le rivendicazioni di sovranità quanto l'incapacità di far rispettare quanto si sia imposto. E in un contesto nel quale l'Europa somma non pochi fattori di debolezza, anche le semplici dichiarazioni, prive di riscontri concreti, possono rivelarsi controproducenti, se non siano percepite come qualcosa di concretizzabile.

L'approccio geopolitico è proprio in questo snodo che interviene: è anche grazie alla geopolitica che si può comprendere, in prima battuta, se e come si possano far seguire a una dichiarazione effetti concreti e, successivamente, che tipo di effetti, immediati e nel medio-lungo periodo, possono disegnarli. Non è poco, anzi è moltissimo, specie in uno scenario globale tanto frammentato e incerto come quello attuale. Ma allora non può na-

²⁶ Il riferimento va al c.d. «modello di Tiebout»: C. Tiebout, *A Pure Theory of Local Expenditures*, in *Journal of Political Economy*, 5, 1956, 416 ss.; su tale modello, v. H.S. Banzhaf – R.P. Walsh, *Do People Vote with Their Feet? An Empirical Test of Tiebout's Mechanism*, in *The American Economic Review*, 3, 2008, 843 ss.

²⁷ Per qualche approfondimento, sul punto specifico, sia consentito il rinvio a P. Passaglia, *Protezione dei dati personali e intelligenza artificiale: il Consiglio d'Europa tra ruolo geopolitico e tutela dei diritti fondamentali*, in *Alexis*, autunno-inverno 2024/2025, 78 ss.

scondersi un certo rammarico derivante dall'elusione di qualunque analisi geopolitica nel momento in cui a livello europeo si rivendicava con forza il proprio essere un attore sovrano nell'ecosistema digitale.

Perché se alle parole debbono seguire i fatti, è bene che i fatti vengano quanto più possibile previsti.

Abstract

La Dichiarazione per la sovranità digitale europea, recentemente approvata dai rappresentanti degli Stati membri dell'Unione europea, è destinata, per i prossimi anni, a orientare la politica dell'Unione in materia di nuove tecnologie. Per quanto la definizione che si è data di «sovranità digitale» esprima, sotto vari punti di vista, una posizione di debolezza dell'Europa, la Dichiarazione appare comunque una rivendicazione di potere che richiede una adeguata concretizzazione. Su questo punto, la mancanza di un approccio geopolitico potrebbe dar luogo a difficoltà in futuro.

The Declaration for European Digital Sovereignty, recently approved by representatives of the Member States of the European Union, is intended to guide EU policy on new technologies for the coming years. Although the definition of “digital sovereignty” reflects, in many ways, Europe’s position of weakness, the Declaration nevertheless appears to be a claim to power that requires adequate implementation. On this point, the lack of a geopolitical approach could give rise to difficulties in the future.

Keywords

Unione europea - sovranità digitale - Dichiarazione per la sovranità digitale europea - ecosistema digitale - geopolitica

L'uso della forza nel dominio cibernetico ai sensi del diritto internazionale*

Rita Caracuzzo

Sommario

1. Premessa: la crescente emergenza delle minacce cibernetiche. -2. Il problema dell'attribuzione (attori statali e non statali). - 2.1 Attribuzione tecnica. - 2.2 Attribuzione legale. - 2.3 Attribuzione politica. - 2.4 L'onere della prova. - 3. La legge applicabile al *cyberspace*: interpretazione dei trattati esistenti o necessità di nuove specifiche regole? - 4. La consuetudine e i principi internazionali. - 4.1 Principio di sovranità e di non intervento. - 4.2 Divieto dell'uso della forza. - 4.3 Il diritto di auto-difesa. - 4.4 Principio di diligenza (*Due diligence*). - 5. La responsabilità internazionale dello Stato. - 5.1 Cessazione delle condotte illegittime e misure riparative. - 5.2 Le contromisure. - 6. Conclusioni e prospettive future. - 6.1 La posizione italiana. - 6.2 La posizione europea.

1. Premessa: la crescente emergenza delle minacce cibernetiche

Le “*Information and Communication Technologies*” (ICT) hanno rivoluzionato il modo in cui gli Stati concepiscono i loro confini ed il modo in cui essi conducono le loro attività. Queste tecnologie hanno sostanzialmente aggiunto una nuova realtà, quella cibernetica, diversa da quella fisica, nella quale gli esseri umani debbono imparare a vivere e gli Stati debbono imparare a relazionarsi. Così, nel c.d. *cyberspace*, sussistono nuove opportunità per lo scambio di informazioni, la comunicazione istantanea e la gestione efficiente dei dati, ma si presentano altresì sfide significative dal punto di vista legale, specialmente quando si tratta dell'uso del dominio cibernetico che gli Stati fanno in ambito internazionale.¹

Questo mondo virtuale (oggi definito come il “quinto dominio”, accanto al territorio, mare, aria e spazio) è infatti creato dal collegamento tra com-

* L'articolo è stato sottoposto, in conformità al regolamento della Rivista, a referaggio anonimo.

¹ Il *cyberspace* fu definito da William Gibson negli anni '80, che lo utilizzò nel suo romanzo di fantascienza “*Neuromante*”, per descrivere un universo digitale accessibile tramite connessioni elettroniche. Esso si riferisce ad uno spazio virtuale creato attraverso l'interconnessione di reti informatiche e sistemi digitali, dove si svolgono comunicazioni, scambio di dati, interazioni sociali e gestione delle informazioni. Alcuni lo definiscono una sorta di “allucinazione consensuale”.

puters, dispositivi abilitati ad Internet, servers, routers ed altri componenti delle infrastrutture di Internet. È proprio Internet la chiave che permette le connessioni tra i vari elementi.²

Ad oggi non esiste una autorità centralizzata che governa l'intero *cyberspace*, ma gli attori che lo popolano sono gli Stati, i Governi, le Organizzazioni intergovernative, le comunità, il settore privato, le organizzazioni della società civile e gli individui.³

Anche le distanze geografiche e le frontiere sono oggi diventate irrilevanti nel contesto cyber, tanto che il c.d. Stato “bersaglio” potrebbe essere colpito attraverso un computer situato dall'altro lato del mondo, in una manciata di secondi.

Per tali motivi la cybersicurezza è considerata oggi uno dei settori chiave per la tutela non solo degli individui ma anche dello Stato in quanto tale

² Va ricordato che ARPANET, la prima rete da cui Internet deriva, era una rete militare finalizzata allo scambio di informazioni. Nel 1958 il governo degli Stati Uniti fondò l'Istituto di ricerca ARPA (*Advanced Research Project Agency*), con l'obiettivo di cercare soluzioni tecniche innovative volte ad assicurare la sicurezza e la disponibilità di una rete di informazioni. Il progetto, sviluppato poi negli anni Sessanta, cercò di costruire una rete di comunicazione militare capace di resistere anche ad un attacco nucleare su larga scala. Con l'avvento dei personal computers negli anni ottanta, gli “utenti” istituzionali e militari cominciarono a rendere partecipi alla rete i membri della comunità scientifica, che iniziarono così a scambiarsi informazioni e dati, nonché a coinvolgere, a loro volta, altri “utenti” comuni. In seguito, nei primi anni degli anni novanta, con i primi tentativi di sfruttamento commerciale, grazie a una serie di servizi offerti da varie aziende, ebbe inizio il vero boom di Arpanet, nel frattempo rinominata Internet, e negli stessi anni nacque una nuova architettura capace di semplificare enormemente la navigazione: il World Wide Web, inventato da Tim Berners-Lee nel 1989.

³ Volendo delineare le caratteristiche del *cyberspace*, potremmo dire che esso ha una portata globale e una natura senza confini; è essenziale per la vita moderna; non è allo stato uno spazio sicuro (sussistono infatti vari tipi di minacce cyber, peraltro sempre più frequenti); è un dominio “contestato”, nel senso che si scontrano e competono in tale ambito diversi interessi e preoccupazioni per la sicurezza; rappresenta un luogo “opaco”, nel senso che gli attori possono facilmente muoversi in anonimato; presenta un forte problema relativo alla “attribuzione” dell'operazione al suo effettivo autore; è “asimmetrico”, nel senso che penetrare all'interno di un network si rivela più semplice che difenderlo dalle minacce dell'aggressione; presenta problemi relativi alla legge applicabile.

La c.d. rivoluzione digitale coinvolge peraltro non solo le infrastrutture civili, ma anche le forze armate, essendo il *cyberspace* utilizzato anche per scopi militari (esistono infatti bombe guidate da GPS satellitari; droni pilotati da remoto o con sistemi di intelligenza artificiale; aerei da combattimento e navi da guerra trasformati in enormi centri di elaborazione dati). Secondo la dottrina del *Network-centric warfare*, anche lo stesso *cyberspace* può essere utilizzato per azioni cibernetiche offensive o di penetrazione delle reti avversarie.

Mentre da un lato si assiste ad una “corsa” degli Stati ad una sempre maggiore digitalizzazione dei propri apparati anche al fine di migliorare l'efficienza e l'efficacia dell'azione amministrativa, giudiziaria e del *welfare*, dall'altro il maggiore affidamento sui sistemi informatici trasforma lo Stato in soggetto maggiormente vulnerabile ad eventuali “attacchi” informatici posti in essere da individui o da altri Stati, idonei a paralizzare il funzionamento dell'intera società.

A ciò si aggiunga un'ulteriore considerazione. Le tecnologie e le competenze cyber sono relativamente semplici ed economiche da acquisire, ragion per cui anche singoli individui (persino giovani teenagers) si sono dimostrati in grado di causare considerevoli danni al Paese, aggredendo ed hackerando un sistema informatico, distruggendo dati presenti sui computers, impedendone il corretto funzionamento, con possibili ripercussioni negative sulla funzionalità delle infrastrutture e conseguente rischio per l'incolumità delle persone.

e costituisce la maggiore preoccupazione dei Governi, quanto meno di quelli dei Paesi più digitalizzati.

La difesa dalla minaccia cibernetica risulta infatti di fondamentale importanza, sia che la minaccia provenga da individui o gruppi criminali e terroristici, sia che essa provenga da altri Stati.⁴

Si discute poi su quali caratteristiche debba avere un “*cyber attack*” per essere considerato un “*armed attack*” ai sensi dell’art. 51 della Carta delle Nazioni Unite o un attacco ricadente nella previsione dell’art. 49 (1) del I Protocollo Addizionale alla Convenzione di Ginevra sulla Protezione delle Vittime di Guerra del 1949, con conseguente applicazione della relativa disciplina (tematica che sarà affrontata in seguito).

Tuttavia, mentre nel caso in cui l’operazione cyber sia commessa da individui o gruppi (criminali o terroristici) o entità private per scopi personali, economici o ideologici, viene in rilievo essenzialmente una questione di legge domestica nazionale (sono quindi le legislazioni dei singoli Stati i cui sistemi informatici vengono aggrediti che troveranno applicazione, in base al principio della sovranità della loro giurisdizione), nel caso in cui invece le attività cyber siano condotte da Stati contro altri Stati, la questione ricade nel mandato del diritto internazionale, che regola appunto le relazioni e i rapporti tra Stati dotati di propria sovranità. Va comunque specificato che, anche se l’operazione cyber risulta commessa da gruppi terroristici, occorrerà sempre verificare se essi agiscano o meno per conto di uno Stato e quindi se l’attività dagli stessi posta in essere, sia attribuibile a quello Stato.

La prima sfida che si presenta è quindi quella di accertare e verificare il soggetto o entità alla quale è attribuibile l’operazione cyber, perché questo comporterà, di conseguenza, l’applicazione di una diversa normativa.

Nel caso poi si accerti che l’attacco cyber provenga da uno Stato, occorre verificare se e fino a che punto le norme del diritto internazionale già esistenti possano e debbano applicarsi al *cyberspace*, ossia occorre verificare quale sia il quadro normativo entro il quale gli Stati dovrebbero operare quando utilizzano le ICT.

⁴ M. Roscini, *Cyber Operations and the Use of Force in International Law*, Oxford, 2014, 16. Le tipologie delle operazioni cyber che possono essere condotte, sono molteplici ma al di là della varia nomenclatura e classificazione occorre preliminarmente distinguere tra “*cyber exploitation*” e “*cyber attack*”. Per “*cyber exploitation*” si intende l’accesso non autorizzato a computers, sistemi informatici o networks, al fine di ottenere informazioni, ma senza compromettere la funzionalità del sistema a cui si accede o la modifica o la cancellazione dei dati ivi residenti. Di contro, i “*cyber attacks*” sono quelle operazioni cyber, sia per offesa sia in difesa, poste in essere al fine di alterare, eliminare, corrompere o negare l’accesso a computer o software con lo scopo di propaganda o inganno; di distruggere parzialmente o totalmente la funzionalità del computer, computer system o network bersaglio, e computers correlati a infrastrutture fisiche; di produrre danni fisici estrinseci ai computers, computer system o network. In ogni caso, anche nel “*cyber exploitation*” vi è una intrusione, accesso non autorizzato e penetrazione nelle reti, ma questo è solo finalizzato a carpire informazioni, mentre il “*cyber attack*” ha finalità distruttive e manipolative dei dati.

2. Il problema dell'attribuzione (attori statali e non statali)

Appare dunque di fondamentale importanza ribadire che uno Stato assume la responsabilità di un atto nel dominio cibernetico sulla base del fatto che l'atto sia allo stesso attribuibile, nel senso che esso sia condotto o comandato da un organo dello Stato, o da una persona o autorità rappresentativa dello Stato. Per contro, non potrà essere attribuita ad uno Stato quella operazione cyber che è condotta da un attore non statale, sia esso un individuo, una organizzazione o persino un gruppo terroristico, a meno che l'attore non statale abbia agito su ordine o secondo le istruzioni, direzione e controllo dello Stato stesso.

L'attribuzione ha diversi aspetti: tecnico, legale e politico.

2.1. Attribuzione tecnica

L'attribuzione tecnica, ossia la determinazione dell'origine territoriale e dell'autore dell'atto cyber, è oggi particolarmente complessa anche perché alcuni attori non-Statali hanno dimostrato capacità informatiche che precedentemente appartenevano solo agli Stati, ragion per cui, a volte, risulta più difficile individuare il responsabile.

Attribuire l'atto o la condotta al suo autore è la base per la contestazione di qualsiasi tipo di responsabilità. Nel contesto cyber, identificare chi si cela dietro la cyber operazione, presenta enormi problemi tecnici.

L'anonimato è infatti una delle più importanti caratteristiche del *cyberspace*.⁵ In ogni caso, le difficoltà nel raggiungere la prova relativa alla identificazione dell'autore dell'attacco cyber non devono certo costituire un deterrente per trattare in maniera approfondita gli aspetti di diritto internazionale collegati.

2.2. Attribuzione legale

Dato per assunto che l'autore dell'operazione cyber sia stato identificato, il problema che sorge è poi di natura legale, ossia quello di stabilire se la condotta realizzata dallo specifico autore possa essere attribuita allo Stato, in base alla legge sulla responsabilità degli Stati. Lo scopo dell'attribuzione consiste quindi in un'operazione normativa volta a stabilire la responsabilità dello Stato che si cela dietro l'atto cyber, individuando il nesso legale tra l'autore dell'atto e lo Stato stesso.

⁵ Internet, in particolare, è un sistema decentralizzato nel quale il protocollo delle comunicazioni divide i dati inviati in molteplici settori che prendono vie e sentieri diversi per poi riassemblarsi una volta giunti a destinazione. È vero che con la cooperazione dei Internet Service Provider (ISP) si può ottenere l'indicazione dell'*IP address*, ossia dell'indirizzo connesso ad Internet attraverso il quale si origina l'operazione, che potrebbe essere associato ad una persona, un gruppo o ad uno Stato. Tuttavia, l'*IP address* potrebbe essere stato falsificato o alterato, o potrebbe rappresentare solo un collegamento intermedio per un aggressore collocato in altro luogo.

Risulta così di fondamentale e cruciale importanza, individuare e capire il motivo dell'attacco cyber e conseguentemente essere in grado di differenziare l'atto criminale (rientrante nell'applicazione della legge nazionale dello Stato aggredito) dall'atto di guerra o di aggressione di uno Stato contro un altro Stato (rientrante nell'ambito del diritto internazionale).

Sebbene non sia astrattamente impossibile che un regime speciale di responsabilità internazionale dello Stato si possa sviluppare come conseguenza delle univoche e peculiari caratteristiche delle cyber operazioni, nella attuale assenza di indicazioni in tal senso, si ritiene che le norme applicabili siano quelle relative alla generale responsabilità degli Stati, contenute nel Capitolo II della Parte Prima delle regole codificate dalla Commissione del Diritto Internazionale, nel Progetto di *Articoli sulla Responsabilità degli Stati per Atti Illeciti Internazionali* (ARSIWA), approvati dall'Assemblea Generale delle Nazioni Unite nel 2001, che riflettono sostanzialmente la legge consuetudinaria internazionale.

Dalla lettura di tali articoli è possibile dedurre che l'attribuzione di una condotta (quindi oggi anche condotta cyber) possa determinarsi in ragione di tre fondamentali criteri.

In primo luogo, si evidenzia il criterio del collegamento istituzionale, che si applica quando la condotta viene automaticamente attribuita ad uno Stato se realizzata da uno dei suoi organi (che esercita funzioni legislative, esecutive, giudiziali od ogni altra funzione), o da individui o entità che esercitano l'autorità di governo.⁶

Il concetto di "organo di Stato" è ampio. Esso ricomprende tutte le persone o entità che hanno tale status secondo la legge nazionale. Così ogni cyber attività effettuata da servizi di intelligence, servizi militari, di sicurezza interna o da altre agenzie statali, comporta una responsabilità dello Stato se viola un obbligo internazionale vincolante per lo Stato stesso.^{7 8} Peraltro, la condotta degli organi dello Stato è attribuibile allo Stato anche

⁶ *Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations*, Cambridge, regole 4 e 5 (manuale prodotto dal lavoro di una equipe di accademici internazionali, indipendenti, esperti di diritto internazionale umanitario e dell'uso della forza).

⁷ M. Roscini, *Cyber Operations and the Use of Force in International Law*, Oxford, 2014, 35. Rientrano nel concetto di organi di Stato, ad esempio, le unità cyber istituite presso alcuni Paesi, come la National Computer Emergency Response Teams (CERTs), che sono responsabili della gestione e del coordinamento della risposta alle emergenze informatiche, come attacchi informatici, vulnerabilità di sicurezza e altre minacce digitali. Il CERT di un Paese è spesso istituito da un ente governativo, con il compito di proteggere le infrastrutture critiche e promuovere la sicurezza informatica a livello nazionale.

⁸ In Italia esiste il Computer Security Incident Response Team – Italia (CSIRT Italia) che è l'equivalente del CERT (oltre ad alcuni CERT ministeriali istituiti in precedenza). Il CSIRT è stato istituito con Decreto del Presidente del Consiglio dei Ministri (DPCM) del 17 febbraio 2017, per coordinare e gestire la risposta agli incidenti di sicurezza informatica a livello nazionale, in conformità con le direttive europee sulla sicurezza cibernetica, come la Direttiva NIS - Network and Information Security del 2016 (direttiva (UE) 2016/1148; è stata adottata il 6 luglio 2016 dal Parlamento e dal Consiglio dell'Unione europea. È entrata in vigore nell'agosto 2016 e l'Italia l'ha recepita con D.Lgs 18 maggio 2018 n. 65; nel 2023 è entrata in vigore la Direttiva NIS 2 (direttiva (UE) 2022/2555 che aggiorna e rafforza le disposizioni della NIS originale. L'Italia l'ha recepita con d.lgs. 138/2024). Il CSIRT italiano è collocato ed operativo presso l'Agenzia per la Cybersicurezza Nazionale (ACN), creata ed istituita dal d.l. 82/2021 (convertito in l. 109/2021) per rafforzare la sicurezza delle infrastrutture critiche italiane e promuovere la resilienza cibernetica.

se essi eccedono nell'esercizio delle funzioni autorizzate o contravvengono alle istruzioni ricevute.

In aggiunta agli atti commessi da organi dello Stato, vi sono gli atti commessi da persone o entità che seppure non propriamente qualificate organi di Stato, sono state tuttavia autorizzate dalla legge nazionale ad esercitare una qualche funzione governativa dello Stato. In questo caso la condotta sarà però attribuibile allo Stato, a condizione che quella persona o entità agisca in tale capacità funzionale (quella autorizzata), nel caso concreto.⁹

In secondo luogo, si valorizza il collegamento fattuale in base al quale la condotta è attribuita ad uno Stato nel caso in cui il soggetto che ha commesso l'illecito stia agendo sotto le istruzioni, direzione o controllo dello Stato stesso (cfr. art.8 ARSIWA e Rule 17 Tallinn Manual 2.0).

La Corte Internazionale di Giustizia ha specificato, in diverse occasioni, che per direzione e controllo si intende il potere dello Stato di far iniziare e cessare l'operazione cyber.¹⁰

Per contro si ritiene che un generico supporto o incoraggiamento dello Stato all'attività dei privati non sia sufficiente per l'integrazione del requisito di attribuzione.

Vi è infine il criterio del "*ex post facto*", o del riconoscimento postumo, in base al quale la condotta viene riconosciuta come propria da uno Stato dopo che essa ha avuto luogo (cfr. art. 11 ARSIWA e Regola 17 Tallinn Manual 2.0).¹¹

2.3. Attribuzione politica

Gli aspetti politici dell'attribuzione si materializzano nell'atto di allocazio-

⁹ Esempi di questo tipo di soggetti includono le società private alle quali è stata concessa l'autorità, dal Governo, per condurre condotte cyber offensive contro un altro Stato, o entità private legalmente autorizzate ad impegnarsi nella raccolta di informazioni informatiche. Se queste corporazioni o società private conducono ad esempio operazioni cyber in difesa di network militari identificati dal Governo, la loro attività sarà attribuibile allo Stato, ma se invece esse eseguono compiti di sicurezza informatica per enti privati, la loro attività non può essere attribuita allo Stato.

¹⁰ M. Roscini, *Cyber Operations and the Use of Force in International Law*, Oxford, 2014, 36. Sono stati considerati esempi di questo tipo: i "Russian Business Network" (RBN), un gruppo di cyber criminali specializzati in *phishing*, *malicious code*, comando e controllo di *botnet*, attacchi DDoS e furti di identità, che fu sospettato di aver eseguito le cyber operazioni contro lo Stato della Georgia, sotto il controllo e per conto della Russia, nell'agosto del 2008; gli "Armenian Cyber Army", un gruppo di hackers che furono accusati dall'Azerbaijan di aver agito sotto la direzione ed il controllo dell'Armenia, nel 2012; un gruppo di patriottici hackers filorussi che l'Estonia indicò come soggetti incoraggiati dalla Russia per condurre l'attacco cyber del 2007.

¹¹ *Il più grande fallimento delle forze speciali americane*, in *ilpost.it*, 24 aprile 2020. La Corte Internazionale di Giustizia, nel caso degli ostaggi di Tehran (*Tehran Hostages case*), ha riconosciuto tale regola espressione del diritto internazionale consuetudinario, attribuendo all'Iran la responsabilità per la detenzione di diversi ostaggi statunitensi tra il 1979 e il 1981 all'interno dell'ambasciata statunitense, che fu occupata. In tal caso il sequestro degli ostaggi, materialmente effettuato da un gruppo di studenti, nel corso della rivoluzione iraniana, era stato approvato dall'ayatollah Khomeini e da altri organi dello Stato iraniano.

ne delle responsabilità. L'attribuzione consiste infatti nell'assegnazione di responsabilità ad uno Stato su base politica, posto che gli Stati non hanno obblighi o aspettative di attribuire le malevoli operazioni cyber a qualcuno. La c.d. fase politica consiste poi nella raccolta di informazioni attraverso canali interstatali, ufficiali e non, per giungere alla individuazione delle potenziali ragioni (spesso geopolitiche) sottese all'attacco e al soggetto, ovvero all'organizzazione da cui esso proviene.

Attribuire la responsabilità politica di un attacco cyber ad un determinato Stato, costituisce una presa di posizione nel panorama internazionale, specie se tale attribuzione avvenga congiuntamente da parte di più Stati.

2.4. L'onere della prova

Il rapporto del 2021 del Gruppo di Esperti Governativi (GGE) costituito nelle Nazioni Unite per affrontare questioni specifiche di rilevanza globale (tra le quali appunto quella della sicurezza informatica internazionale), afferma che le accuse relative alla commissione di atti illeciti mosse contro gli Stati, devono essere provate.

Nell'ambito del *cyberspace*, oltre alle consuete difficoltà probatorie, si aggiunge l'ulteriore difficoltà dovuta al fatto che normalmente il luogo da dove si origina l'attacco cyber è situato al di fuori del territorio dello Stato vittima. Questo rende ancora più complicato acquisire gli elementi di prova.

Le tematiche maggiormente discusse in tale ambito riguardano sia l'individuazione del soggetto a cui spetta l'onere della prova, sia la determinazione del grado di prova che può essere ritenuto sufficiente per l'affermazione della responsabilità.

Solitamente spetta a colui che lamenta il danno, provarlo.

Al contrario, in ambito cyber, potrebbe ipotizzarsi, per alcuni, una inversione dell'onere della prova, come già accade ad esempio nei casi di tortura o trattamenti inumani o degradanti. Nella giurisprudenza della Corte europea dei diritti dell'uomo (Corte EDU), infatti, l'onere della prova in tali casi, ai sensi dell'art. 3 della Convenzione europea dei diritti dell'uomo (CEDU), segue il principio del "*burden-sharing*", ossia del criterio secondo cui, se una persona presenta una denuncia credibile e sufficientemente dettagliata (ad esempio sostenendo di aver subito tortura mentre era sotto la custodia dello Stato), spetta allo Stato dimostrare che le accuse sono infondate. Pertanto, anche in ambito cyber, si sostiene, la tematica è ancora oggetto di discussione.

Se il soggetto presenta una denuncia specifica e completa, sarà lo Stato a provare l'infondatezza delle accuse.

La legge internazionale non prescrive uno standard generale di prova per tutti gli atti illeciti, tanto che le Corti e i tribunali internazionali hanno elaborato un proprio standard probatorio specifico per ciascun caso, non sempre uniforme. In linea generale può comunque affermarsi che il grado della prova richiesto sarà tanto più alto quanto più grave è l'accusa da provare.¹²

¹² M. Roscini, *Cyber Operations and the Use of Force in International Law*, Oxford, 2014, 98.

Nonostante la sua cruciale importanza, il Tallinn Manual 2.0 non affronta in profondità lo standard della prova richiesto nel cotesto cyber; l'unico riferimento, peraltro superficiale, è contenuto nel Commentario di alcune regole (7, 15 e 82), che comunque non stabiliscono un criterio definitivo e lasciano spazio all'interpretazione e alla prassi statale.¹³

In ogni caso, va dato atto che, in un contesto di autodifesa, la realtà è che gli Stati debbano prendere le proprie determinazioni rispetto all'attribuzione dell'operazione cyber ad un altro Stato, molto rapidamente, prima di rispondere, senza quindi avere tempo per acquisire tutte le dovute e necessarie informazioni che sono invece spesso disponibili in un contesto non cyber.¹⁴

3. La legge applicabile al cyberspace: interpretazione dei trattati esistenti o necessità di nuove specifiche regole?

Nel dominio cibernetico gli Stati hanno raggiunto importanti accordi solo per quanto attiene la disciplina dei cybercrime. Oltre alla Convenzione di Budapest del 2001 del Consiglio d'Europa sulla criminalità informatica, è di recente approvazione la Convenzione delle Nazioni Unite, adottata il 24 dicembre 2024 dall'Assemblea Generale dell'ONU¹⁵. Essa rappresenta un passo significativo nella lotta globale contro i crimini informatici, mirando non solo a prevenire e combattere il cybercrime in modo più efficiente, ma rafforzando altresì la cooperazione internazionale e fornendo assistenza tecnica agli Stati, in particolar modo a quelli in via di sviluppo.

Nel settore del cyberspazio in generale, non ci sono norme pattizie che

¹³ Alcuni ritengono che un minimo livello di azione in difesa, contro un attacco cyber, può essere effettuato anche quando non vi è certezza di chi è stato l'effettivo responsabile. Sarà rilevante, in tal caso, valorizzare alcune circostanze quali: lo stato delle relazioni internazionali tra i due Paesi (il probabile aggressore e lo Stato vittima), i precedenti coinvolgimenti dello Stato sospettato in attacchi cyber, la natura del sistema attaccato, la natura ed il grado di sofisticazione dei metodi e dei mezzi utilizzati, il comportamento dello Stato sospettato nel periodo successivo all'attacco. Assistiamo quindi ad un approccio di standard di prova ben lontano dal grado di prova "chiara e convincente". Tuttavia, altri criticano questa impostazione sottolineando che lo standard di prova nel contesto cyber non dovrebbe essere più basso solo perché risulta di difficile raggiungimento. È di fondamentale importanza evitare false attribuzioni, penalizzando così lo Stato sospettato ed iniziare un processo di attacchi cyber in difesa che innescano ulteriori operazioni cyber. Risulta pertanto più appropriato il criterio dello standard di prova "chiara e convincente". Per contro, il criterio della prova "oltre ogni ragionevole dubbio" si rivelerebbe irrealistico.

La discussione rimane comunque aperta in ambito internazionale, con alcuni Stati che preferiscono standard più elevati prima di formulare accuse ufficiali.

¹⁴ *Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations*, Cambridge. Il gruppo di esperti, autori del Tallinn Manual 2.0, ritengono che gli Stati dovrebbero agire ragionevolmente, come farebbero in circostanze identiche o simili, nel rispondere a tali situazioni. La ragionevolezza poi, dipende da diversi fattori, quali l'affidabilità delle informazioni, le dimensioni dell'attacco, la sua natura e la specificità dei diritti violati coinvolti.

¹⁵ Documento A/79/460 del 27 novembre 2024 della United Nations General Assembly *Countering the use of information and communications technologies for criminal purposes*

disciplinano espressamente le azioni degli Stati e di altri attori, ma questo non significa però che esso debba restare senza disciplina. Tuttavia, mentre alcuni Stati ritengono che l'attuale quadro di diritto internazionale sia adattabile al *cyberspace*, altri chiedono alla comunità internazionale di negoziare un trattato multilaterale vincolante che possa definire con maggiore precisione delle regole che tengano conto della unicità e complessità del dominio cibernetico.

Importante riferimento è costituito altresì dal diritto consuetudinario internazionale, le cui norme sono in continua evoluzione.

Preliminarmente, in ambito cyber, occorre distinguere tra lo *jus ad bellum* e lo *jus in bellum*. Lo *jus ad bellum* si riferisce alle condizioni che giustificano il ricorso alla guerra o comunque all'uso della forza da parte di uno Stato. I principi fondamentali del *jus ad bellum* sono sanciti nella Carta delle Nazioni Unite. Lo *jus in bellum* disciplina invece le regole che disciplinano la condotta dei belligeranti durante il conflitto armato, indipendentemente dalla sua legittimità iniziale. È noto anche come Diritto Internazionale Umanitario (DIU) ed ha lo scopo di limitare gli effetti della guerra proteggendo i civili e i combattenti che non partecipano più alle ostilità.

Nella presente trattazione verrà affrontato solo l'aspetto dello *jus ad bellum* (uso della forza da parte di uno Stato nei confronti di altro Stato), nel quale un ruolo fondamentale, viene riconosciuto alla Carta delle Nazioni Unite, ossia il trattato istitutivo dell'Organizzazione delle Nazioni Unite (ONU), firmato il 26 giugno 1945 a San Francisco.

Gli esperti si sono quindi chiesti se ed in quali termini i principi ivi elencati siano applicabili al *cyberspace*.

4. La consuetudine e i principi internazionali

4.1. Principio di sovranità e di non intervento

Tra i principi fondamentali di diritto internazionale, vi è senz'altro il principio di sovranità, riconosciuto dall'art. 2, par. 1, della Carta delle Nazioni Unite e a livello consuetudinario (c.d. principio generale di "*superiorem non recognoscens*"). In virtù di tale principio, gli Stati hanno il diritto di scegliere liberamente il proprio sistema politico, sociale, economico e culturale, nonché esercitare il controllo sovrano sulle risorse e sulle attività che avvengono nel proprio territorio. Esso rappresenta l'indipendenza di uno Stato, ossia il diritto di esercitare le funzioni dello Stato con esclusione di ogni altro Stato. Tutti gli Stati sono giuridicamente uguali, senza supremazia dell'uno nei confronti dell'altro. Nel corso degli ultimi anni è in corso un intenso dibattito sull'applicazione di tale principio al *cyberspace*. Nel contesto delle ICT, questo potrebbe significare che gli Stati hanno il diritto di regolare e controllare le infrastrutture ICT all'interno dei propri confini, nonché di adottare tutte le misure necessarie per proteggere la sicurezza nazionale e la privacy dei loro cittadini.

Ai sensi della Regola n. 1 del Tallinn Manuale 2.0, il principio di sovranità si applica al *cyberspace*. Nelle regole seguenti viene altresì specificato che lo Stato esercita la sua sovranità, controllo e protezione sulle strutture cyber

siano esse pubbliche o private, e che i trattati internazionali possono operare restrizioni sull'esercizio della sovranità nel territorio dello Stato, ma ciò dipende da una scelta dello Stato stesso in ordine alla sua adesione o meno al trattato internazionale.

Corollario del principio di sovranità è il principio di non intervento, in base al quale gli Stati devono astenersi dall'intervenire negli affari interni ed esterni degli altri Stati. Si ritiene che anche questo principio possa ritenersi applicabile al *cyberspace*, ragion per cui gli Stati non devono porre in essere attacchi informatici o comunque interferire nei sistemi informatici di altri Stati che possano minacciare o mettere in pericolo la sovranità di un altro Stato.

Il principio del non intervento viene invocato nella misura in cui l'attacco cyber non assurge al livello dell'"attacco armato" che implica "la minaccia o l'uso della forza" ex art. 2, par. 4, della Carta delle Nazioni Unite (principio che verrà affrontato nel prossimo paragrafo). Se sussiste il superamento di tale soglia, non si invoca più la violazione del principio di non intervento, ma appunto la violazione della Carta delle Nazioni Unite, che comporta poi una serie di conseguenze, quali la possibilità di esercitare il diritto di difesa.

Si ritiene che il principio di non intervento sia espressione del diritto internazionale consuetudinario. Anche se esso non è espressamente menzionato nella Carta delle Nazioni Unite, è stato incorporato in molti accordi internazionali, tra cui la Dichiarazione sulle Relazioni Amichevoli tra Stati del 1970 adottata dall'Assemblea Generale dell'ONU con la Risoluzione 2625¹⁶.

4.2. Divieto dell'uso della forza

Altro principio di diritto internazionale, la cui applicazione nel *cyberspace* è discussa, è quello relativo al divieto dell'uso della forza, previsto dall'art. 2, par. 4, e dall'art. 51 della Carta delle Nazioni Unite ed ormai accettato come regola di diritto internazionale consuetudinario (e quindi applicabile anche a Stati che non sono membri delle Nazioni Unite). In virtù di tale principio ogni Stato deve astenersi dal minacciare o usare la forza contro l'integrità territoriale o l'indipendenza politica di un altro Stato. Ambientato nel *cyberspace*, questo significherebbe proibire ad ogni Stato di "usare

¹⁶ L'intervento proibito si compone di due elementi principali: deve essere di natura coercitiva; e deve riguardare questioni in cui allo Stato è consentito, sulla base del principio di sovranità statale, di decidere liberamente, includendo sia gli affari interni che gli affari esterni. In particolare, la natura coercitiva consiste nel costringere lo Stato a fare o ad omettere di fare qualcosa all'interno della sfera del proprio dominio, nonché a restringere o privare il controllo dello Stato su una materia o un settore protetto.

La Carta delle Nazioni Unite in realtà, seppur non codifica il principio di non intervento di uno singolo Stato nei confronti di un altro, afferma, all'art. 2, par. 7, che le Nazioni Unite non sono autorizzate ad intervenire in questioni che appartengono sostanzialmente alla competenza interna di uno Stato. Vi è quindi, in ogni caso, il riconoscimento del c.d. "dominio interno" di uno Stato, ossia di un insieme di beni e risorse che appartengono allo Stato e che sono situati all'interno del suo territorio. Si riflette sul fatto che questo concetto possa ricomprendere anche le infrastrutture digitali e le risorse informatiche, e quindi applicarsi al dominio cibernetico.

la forza” anche tramite l'utilizzo delle strutture ICT, ossia utilizzando il computer come “arma”.

Le condizioni che debbono verificarsi, affinché possa operare l'art. 2, par. 4, sono le seguenti: in primo luogo, come sopra specificato, la cyber operazione deve essere attribuibile allo Stato (e non ad un cittadino privato o ad un gruppo criminale); in secondo luogo la cyber operazione deve consistere nella “minaccia” o nell’“uso della forza”; ed infine la minaccia o l'uso della forza devono essere esercitati nell'ambito delle condotte e relazioni internazionali.

Il riferimento alle “relazioni internazionali” nell'art. 2, par. 4, sta a significare che la minaccia o l'uso della forza debbono non solo provenire da uno Stato, ma debbono altresì essere diretti nei confronti di un altro Stato. Pertanto, non è proibito agli Stati minacciare od utilizzare la forza ricorrendo ad operazioni cyber, contro individui o gruppi criminali, purché tali operazioni non influenzino l'integrità territoriale o l'indipendenza politica di un altro Stato¹⁷.

4.3 Il diritto di autodifesa

Il diritto di autodifesa è disciplinato dall'art. 51 della Carta delle Nazioni Unite, in base al quale «nessuna disposizione del presente Statuto pregiudica il diritto naturale di autotutela individuale o collettiva, nel caso che abbia luogo un attacco armato contro un Membro delle Nazioni Unite, fintantoché il Consiglio di Sicurezza non abbia preso le misure necessarie per mantenere la pace e la sicurezza internazionale. Le misure prese dai Membri nell'esercizio di questo diritto di autotutela sono immediatamente portate a conoscenza del Consiglio di Sicurezza e non pregiudicano in alcun modo il potere e il compito spettanti, secondo il presente Statuto, al Consiglio di Sicurezza, di intraprendere in qualsiasi momento quell'azione che esso ritenga necessaria per mantenere o ristabilire la pace e la sicurezza internazionale».

Da questa disciplina possiamo dedurre le principali caratteristiche del diritto di autodifesa. Esso, in primo luogo, costituisce la risposta ad un “attacco armato”, la cui scala ed effetti sono stati esaminati nel paragrafo precedente. Vi è pertanto la preoccupazione che l'attribuzione di un significato troppo ampio al concetto di divieto dell'uso della forza, possa poi legittimare oltre misura l'esercizio del diritto di autodifesa posto in essere dallo Stato vittima, con conseguente instabilità dei rapporti internazionali. Esso deve essere poi esercitato contro lo Stato responsabile dell'attacco. Da questo si desume che uno Stato non può esercitare il proprio diritto di autodifesa contro un altro Stato se l'attacco armato è stato effettuato, non

¹⁷ Per quanto riguarda “se e quando” una operazione cyber integra il requisito della “minaccia” o dell’“uso della forza”, la Regola 69 del Tallinn Manual 2.0 prevede che essa è tale quando la sua portata e i suoi effetti sono paragonabili a operazioni non informatiche che raggiungono il livello dell'uso della forza. Si assiste quindi ad una perfetta equiparazione tra operazioni cyber e non. Rimane tuttavia aperta la questione del livello di portata ed effetti che deve essere raggiunto al fine di qualificare un atto come “uso della forza”. Infatti, l'art. 2, par. 4, non definisce il concetto di “forza”.

da quello Stato, ma da individui o gruppi non statali (a meno che, come specificato nel capitolo relativo all'attribuzione, gli individui e i gruppi non statali operavano in nome o per conto dello Stato). Occorre notificare il Consiglio di Sicurezza dell'ONU, e consiste nell'adottare misure individuali o collettive.

In ogni caso, la risposta in autodifesa contro un'operazione cyber che integra il requisito dell'attacco armato, deve incontrare i requisiti della necessità, della proporzionalità e dell'immediatezza.¹⁸

Per quanto concerne la tipologia delle risposte possibili, invocando il diritto di autodifesa, esse possono essere individuali (ossia da parte del solo Stato vittima), che collettive. Queste ultime sono disciplinate dall'art. 51 della Carta delle Nazioni Unite, dal diritto internazionale consuetudinario e, nell'ambito del dominio cibernetico, sono prese in esame nella Regola 74 del Tallinn Manual 2.0. Esse consistono nel condurre una difesa congiunta insieme ad altro Stato che presta assistenza allo Stato vittima dell'attacco armato. Tuttavia, come specificato dalla Corte Internazionale di Giustizia nel caso *Nicaragua* (Nicaragua Vs. Stati Uniti, sentenza CIG del 1986), prima che uno Stato possa prestare assistenza ad un altro Stato in autodifesa collettiva, deve aver ricevuto una specifica richiesta in tal senso dallo Stato vittima, e deve operare nei limiti della richiesta e del consenso di quest'ultimo. L'autodifesa collettiva può essere esercitata sulla base di un trattato precedentemente concluso o di un accordo *ad hoc*.¹⁹

Infine, l'art. 51 della Carta delle Nazioni Unite richiede che l'adozione di misure individuali o collettive in autodifesa, debbano essere segnalate immediatamente al Consiglio di Sicurezza delle Nazioni Unite. In caso di attacco informatico potrebbe essere difficile rispettare tale prescrizione, considerate le caratteristiche intrinseche e la natura del *cyberspace*. Ci si è

¹⁸ Anche se l'art. 51 della Carta delle Nazioni Unite non menziona queste caratteristiche, nel *Nuclear Weapons Advisory Opinion* (1993) la Corte Internazionale di Giustizia ha ribadito che costituisce regola del diritto consuetudinario internazionale, il fatto che l'esercizio del diritto di autodifesa rispetti le condizioni della necessità e proporzionalità. Tale regola, si aggiunge, si applica anche all'art. 51 della Carta delle Nazioni Unite, qualunque sia il mezzo di forza utilizzato. *Legality of the Use by a State of Nuclear Weapons in Armed Conflict*, in CIG, in icjci.org.

¹⁹ Ad esempio, gli alleati della NATO (North Atlantic Treaty Organization firmato a Washington, D.C. il 4 aprile 1949), hanno concordato che, all'art. 5, «un attacco armato contro uno o più di loro in Europa o Nord America sarà considerato un attacco contro tutti loro e di conseguenza concordano che, se si verifica un tale attacco armato, ciascuno di loro, nell'esercizio del diritto all'autodifesa individuale o collettiva riconosciuto dall'articolo 51 della Carta delle Nazioni Unite, assisterà la Parte o le Parti così attaccate». In ambito della cyber difesa, già dal 2008 con l'adozione della prima «Politica sulla Cyber Difesa», dopo gli attacchi informatici contro l'Estonia, la NATO ha provveduto ad estendere tale concetto anche sul piano informatico, fondando altresì un centro di eccellenza per la ricerca e lo sviluppo di strategie di cyber difesa, il Cooperative Cyber Defence Centre of Excellence (CCDCOE). Un altro esempio, ma di accordo ad hoc, è quello che è stato alla base dell'assistenza fornita al Kuwait da una coalizione di Stati nel 1990, in risposta all'attacco armato dell'Iraq. Nu merosi in tale settore sono altresì i Memoranda of Understanding (MoU) tra i vari Stati. In ambito cyber, le operazioni in autodifesa sono quindi ammissibili quando impiegate da forze che operano ai sensi di tali accordi, fatti salvi i requisiti di necessità, proporzionalità e immediatezza che si applicano non solo alle condotte individuali ma anche a quelle collettive.

chiesti quindi, quali potrebbero essere le conseguenze della mancata segnalazione al Consiglio di Sicurezza. Sul punto, la Corte Internazionale di Giustizia, nel caso *Nicaragua* sopra richiamato, ha affermato che tale dovere di segnalazione non riflette il diritto internazionale consuetudinario, e potrebbe costituire uno dei fattori che indicano se lo Stato in questione era convinto o meno di agire in autodifesa. Il fatto cioè che non effettui la prescritta segnalazione al Consiglio di Sicurezza relativa all'azione di autodifesa, potrebbe essere interpretato come mancanza di convinzione sull'attacco armato ricevuto o sull'individuazione del responsabile. La mancata comunicazione, tuttavia, non influenza di per sé la legalità dell'atto di autodifesa, che rimane quindi legittimo purché rispetti tutti i requisiti richiesti di necessità, proporzionalità e immediatezza.

4.4 Principio di diligenza

Altro importante principio di cui si discute l'applicazione al *cyberspace* è quello del *Due diligence* (o "*obligation of vigilance*"), ossia il principio in base al quale ciascuno Stato deve esercitare il dovere di diligenza per non permettere o porre fine alle operazioni informatiche ostili (quindi non solo ad attacchi armati ma a tutte le operazioni cyber illegittime) poste in essere nel proprio territorio da un altro Stato e dirette ad uno Stato terzo. Per tale ragione vi sarebbe quindi l'obbligo, per ciascuno Stato, di adottare ogni ragionevole misura al fine di prevenire, eliminare o mitigare il potenziale e significativo rischio che il proprio territorio, ossia le proprie infrastrutture informatiche, anche se collocate all'estero purché sotto il proprio controllo, siano utilizzate per compiere atti illegittimi, colpire e danneggiare Stati terzi.

Mentre per alcuni l'applicazione di tale principio sarebbe obbligatoria, per altri è opportuna. Appartiene alla prima tesi la posizione degli esperti del Tallinn Manual 2.0, che disciplina il dovere di diligenza nelle Regole 6 ss. Affermano gli esperti che tale principio è basato sul diritto internazionale generale, e deriva dalla nozione stessa di "sovranità", richiedendo che lo Stato protegga all'interno del proprio territorio, anche il diritto degli altri Stati.

Esso è vincolante anche nel contesto del *cyberspace*, posto che le nuove tecnologie debbono essere soggette al diritto internazionale preesistente. Il "*due diligence*" presuppone il coinvolgimento minimo di tre parti: lo Stato bersaglio dell'operazione cyber; lo Stato territoriale che è soggetto al dovere di diligenza; e l'autore dell'operazione cyber, indipendentemente che esso sia una persona, una società, un gruppo non statale o un altro Stato. Occorre infatti evidenziare che la parte che lancia l'operazione cyber, può operare da remoto, ossia dal territorio di uno Stato terzo. Si pensi ad esempio ad un gruppo hacker collocato nello Stato "A" che porta avanti un attacco cyber contro lo Stato "B", usando strutture collocate nello Stato "C". In tal caso se lo Stato "C", a conoscenza dell'utilizzo improprio delle sue strutture, non prende misure fattibili per mettere fine a tale operazione, viola il principio della dovuta diligenza.

Uno Stato deve altresì rispettare il dovere di diligenza su infrastrutture

cyber governative proprie, pur se allocate all'estero (si pensi ad esempio ad una installazione militare in un paese straniero; infrastrutture cyber su piattaforme in alto mare o nello spazio internazionale o nelle sedi diplomatiche).

Appare in ogni caso pacifico che, il danno materiale che lo Stato bersaglio subisce, deve derivare da una operazione cyber che costituisce un atto illecito internazionale, ossia che sia contraria al diritto e in violazione degli obblighi internazionali. Tale danno deve altresì provocare conseguenze gravemente negative ("*serious adverse consequences*"). Si pensi ad esempio al caso in cui lo Stato "A" lancia un "*malware*" che è controllato e coinvolge la infrastruttura allocata nello Stato "B". Lo Stato "B" è consapevole di tale operazione. Il "*malware*" manipola il funzionamento del sistema di controllo di un gasdotto nello Stato "C", causando un'esplosione. Lo Stato "B" deve prendere misure per far immediatamente cessare l'operazione, se ha la possibilità e capacità di farlo. Si applica in tal caso il dovere di diligenza. La conoscenza è altro elemento costitutivo per l'applicazione di questo principio. Esso si applica infatti solo se lo Stato ha conoscenza che il suo territorio viene usato per un atto ostile cyber contro un altro Stato. Sicuramente risulta estremamente difficile per lo Stato colpito, dimostrare che lo Stato dal quale è partito l'attacco cyber, ne era a conoscenza, ma questo, affermano gli esperti, non rende la regola discutibile.

5. Responsabilità internazionale dello Stato

Il principio di non intervento ed il divieto dell'uso della forza (disciplinato dall'art. 2, par. 4, della Carta delle Nazioni Unite e dal diritto consuetudinario), con conseguente possibilità di esercitare il diritto di autodifesa (previsto dall'art. 51 della Carta delle Nazioni Unite), trattati nei capitoli precedenti, implicano, come detto, che vi sia un "attacco armato", ossia che l'uso della forza superi una determinata soglia, in termini di danni materiali alla proprietà, perdita di vite, lesioni alle persone o grave disfunzionamento delle infrastrutture critiche.

Gli attacchi cyber che non causano tali conseguenze possono tuttavia costituire illeciti atti di intervento negli affari interni di altro Stato. In tale settore si invoca il principio della responsabilità dello Stato, ritenuto applicabile anche al *cyberspace*. Del resto, si afferma, assicurare che uno Stato sia responsabile per la violazione della legge internazionale, contribuisce a mantenere ordine e sicurezza internazionale. E questo vale anche per il mondo digitale. In tali casi gli Stati, in autotutela, possono adottare contromisure non coercitive (inclusi attacchi informatici al di sotto del livello di uso della forza) e ricorrere al Consiglio di Sicurezza delle Nazioni Unite. Tra le codificazioni del principio di responsabilità degli Stati nel *cyberspace* e della sua disciplina, molto importante è quella del 2001, ossia il progetto *ARSIWA* (*Articles on Responsibility of States for internationally wrongful acts*), riconosciuto e adottato dall'Assemblea Generale delle Nazioni Unite. Questo atto non costituisce un trattato, ma codifica il diritto internazionale generale, quindi, contiene norme che sono considerate vincolanti sulla base del diritto consuetudinario internazionale, con la conseguenza che esse

influenzano ed orientano le decisioni dei giudici nelle corti e nei tribunali internazionali.²⁰ Spesso, infatti, tali principi sono stati utilizzati dalla Corte Internazionale di Giustizia nelle proprie decisioni.

Gli ARSIWA sono quindi un insieme di principi che stabiliscono le regole fondamentali sulla responsabilità degli Stati quando commettono atti contrari al diritto internazionale. Tale atto si basa su tre concetti chiave: in primo luogo si afferma che uno Stato è responsabile quando viola un obbligo internazionale e la violazione è ad esso attribuibile; lo Stato responsabile deve poi cessare l'atto illecito e riparare il danno; ed infine i mezzi di riparazione includono la restituzione, il risarcimento e la soddisfazione. La finalità è quella di cercare di ripristinare e mantenere l'ordine e la pace nella società internazionale dopo che è stata momentaneamente turbata dal verificarsi di un atto illecito di uno Stato. Il ripristino del comportamento dello Stato, alla conformità con il diritto internazionale, è di vitale importanza per la credibilità del diritto e quindi per la pace e sicurezza delle Nazioni.

Sussiste ed integra un atto illecito internazionale di uno Stato quella condotta consistente in un'azione od omissione che è attribuibile a quello Stato e che costituisce una violazione di un obbligo internazionale dello Stato, secondo le norme del diritto internazionale (e non secondo le norme del diritto interno nazionale).

Ai sensi della Regola 19 del Tallinn Manual 2.0 e del Capitolo V della Parte I delle ARSIWA, l'illiceità di un atto che implichi operazioni informatiche è esclusa nel caso di consenso, legittima difesa, contromisure, necessità, forza maggiore o pericolo. Qualora quindi sussista una di queste circostanze, l'azione o omissione in questione non sarà "illecita" e pertanto, lo Stato che adotta la condotta richiesta o la omette non avrà la responsabilità per quella che altrimenti sarebbe una violazione illecita di un obbligo nei confronti dello Stato danneggiato.²¹

²⁰ Dipartimento federale degli affari esteri DFAE, *Le 11 norme per un comportamento responsabile*, in *eda.admin.ch*, 7 aprile 2021

²¹ Per quanto riguarda il "consenso", questo preclude che l'operazione informatica portata a conoscenza dell'altro Stato, prima di effettuarla, possa essere ritenuta poi un atto illecito. Si pensi, ad esempio, ad uno Stato che acconsente che un altro Stato assuma temporaneamente il controllo di determinati elementi della sua infrastruttura informatica per identificare e rispondere ad attività malevole che ivi si verifichino. Qualora ciò accada, lo Stato non può poi lamentare una violazione dell'obbligo di sovranità o del principio del non intervento da parte dello Stato che ha assunto il controllo dell'infrastruttura informatica. Ovviamente il consenso deve essere stato dato liberamente, ossia non deve essere a sua volta il risultato di una minaccia o coercizione. Non si ritiene invece sia necessario un consenso in forma espressa, posto che anche un consenso implicito può ritenersi sufficiente.

L'illiceità dell'atto internazionale è altresì esclusa quando una operazione cyber si qualifica come autodifesa contro un attacco armato, o si qualifichi come una contromisura nei confronti di un precedente atto illecito subito. La contromisura consiste infatti nelle azioni o strategie adottate per mitigare i danni, ripristinare i sistemi compromessi e prevenire futuri attacchi cyber, nel senso che vi è una sostanziale interruzione degli obblighi di rispetto del diritto internazionale da chi riceve l'offesa, purché siano rispettati i requisiti di proporzionalità, temporaneità e reversibilità (concetti che si svilupperanno più avanti). Altra causa di esclusione del significato di un atto illecito è costituita dalle situazioni di "forza maggiore". Il termine si riferisce a circostanze che implicano il verificarsi di una forza irresistibile o di un evento imprevisto, al di fuori del controllo dello Stato, che rende

5.1. Cessazione delle condotte illegittime e misure riparative

Una volta riconosciuta la responsabilità di uno Stato per aver posto in essere un atto internazionale illegittimo (sia esso anche un atto cyber), sorgono in capo a quello Stato delle obbligazioni c.d. secondarie, *ex art. 30 delle ARSIWA*, tra le quali quella di porre fine alla condotta lesiva se si tratta di una condotta permanente (azione positiva), e quella di offrire adeguate assicurazioni e garanzie di non ripetizione della condotta stessa (condotta negativa).

In tal senso si esprime anche la Regola 27 del Tallinn Manual 2.0 secondo la quale «uno Stato responsabile deve porre fine a un atto illecito a livello internazionale commesso con mezzi informativi e, se del caso, fornire assicurazioni e garanzie di non ripetizione».

Si discute sulla forma e consistenza di tali assicurazioni e garanzie. Ci si chiede in particolare se esse possano essere effettuate solo in forma verbale (come, ad esempio, una dichiarazione di riconoscimento delle responsabilità per aver violato il diritto internazionale, un formale atto di scuse, una manifestazione di dispiacere o comunque altre forme idonee a fornire una “soddisfazione” per lo Stato vittima, come ad esempio l’omaggio alla bandiera o semplici dichiarazioni orali nel contesto di un discorso ufficiale) o siano invece necessari atti concreti e più efficaci.

Lo Stato ha altresì, *ex art. 31 delle ARSIWA*, l’obbligo di riparare integralmente il pregiudizio, sia materiale che morale, causato dall’atto internazionale illecito. Questo può esplicarsi sia attraverso il ripristino della situazione quo ante, ossia quella che esisteva prima che venisse commesso l’atto illegittimo, oppure, quando ciò non è possibile, attraverso una compensazione che copra tutti i danni economici subiti.

5.2 Le contromisure

In linea generale, uno Stato vittima può adottare delle contromisure, ossia delle azioni od omissioni dirette contro lo Stato che si sia reso responsabile di un atto internazionale illecito nei suoi confronti, al fine di indurre quello Stato a conformarsi ai propri obblighi. Le contromisure, per essere considerate lecite, possono quindi essere prese solo in risposta e successivamente al verificarsi di un precedente atto illecito. Non possono essere adottate per scopi preventivi. Le contromisure possono consistere in atti illeciti, come ad esempio la sospensione del rispetto di un trattato, purché questa sia limitata al convincere lo Stato che ha precedentemente violato gli obblighi, a cessare tale violazione. Di fatto le contromisure consistono in una sospensione dell’obbligo del rispetto del diritto internazionale, accordata allo Stato vittima, purché ci sia il rispetto dei requisiti della proporzionalità, temporaneità e reversibilità.

Esse sono disciplinate dagli artt. 49 ss. delle ARSIWA ed analizzate nel Tallinn Manual 2.0.

materialmente impossibile nelle circostanze, adempiere all’obbligo internazionale.

Lo scopo delle contromisure, quindi, sia cyber che non, è quello di indurre lo Stato responsabile dell'atto illecito, a rispettare gli obblighi giuridici che gravano su esso, nei confronti dello Stato vittima. Il fine è quindi quello di far sì che lo Stato responsabile cessi la sua azione od omissione illecita e provveda a fornire allo Stato vittima assicurazioni e garanzie rispetto al non verificarsi di un nuovo atto illecito, e risarcisca altresì il dovuto (Regole 27 e 28 del Tallinn Manual e art. 49 delle ARSIWA). Esse sono quindi concepite come un rimedio ideato per ripristinare relazioni legittime tra gli Stati interessati, ed è per tale motivo che non sono ammissibili nel caso in cui l'atto illecito sia terminato, è altamente improbabile che si ripeta e lo Stato responsabile abbia già fornito assicurazioni, garanzie e riparazione del danno. Punizione e ritorsione non sono scopi ammissibili, non avendo le contromisure natura punitiva.

Inoltre, sempre secondo le ARSIWA, le contromisure non possono essere prese (e se prese debbono essere sospese), se la controversia in questione è pendente dinanzi ad una corte o tribunale, che può emettere una decisione vincolante in materia e l'attività illecita è cessata. Sul punto gli esperti internazionali hanno specificato che l'espressione "corte o tribunale" deve riferirsi a «qualsiasi procedura di risoluzione delle controversie di terze parti, qualunque sia la sua designazione». La corte o il tribunale devono avere il potere di ordinare le misure provvisorie in grado di proteggere lo Stato presumibilmente danneggiato. Si è altresì concordato che, anche se l'atto illecito, pur non essendo del tutto cessato, non causa più danni ulteriori e le negoziazioni in corso tra le parti sono in buona fede, le contromisure sono vietate²².

Per quanto riguarda chi sia titolato a prendere le contromisure, l'art. 49 delle ARSIWA, stabilisce che esse possono essere adottate dallo "Stato

²² Per quanto riguarda le loro caratteristiche, le contromisure debbono essere temporanee (ossia possono essere adottate fino a quando lo Stato alle quali sono dirette, non si conforma ai propri obblighi); proporzionali al pregiudizio subito; e comunque tali da non pregiudicare la tutela dei diritti umani fondamentali e le altre norme imperative del diritto internazionale generale (art. 50, 51 e 53 delle ARSIWA). Debbono inoltre essere reversibili, nel senso che devono essere sospese o revocate una volta che lo Stato responsabile dell'illecito internazionale cessa la sua condotta illecita e ripristina la conformità del suo comportamento al diritto internazionale. Pertanto, le contromisure non devono causare danni permanenti, né creare situazioni irreversibili (ad esempio, un embargo commerciale potrebbe essere revocato una volta che lo Stato bersaglio cessa il comportamento illecito, mentre la distruzione irreparabile di infrastrutture civili non sarebbe una contromisura legittima).

Nell'applicazione di tali regole nell'ambito cyber, si discute sul significato e la portata del concetto di "proporzione" rispetto al danno subito. Si ritiene in particolare che la "proporzione" vada commisurata all'entità del danno sofferto, alla gravità dell'atto illecito posto in essere, alla tipologia dei diritti che sono stati violati nello Stato vittima ed alla necessità di indurre lo Stato responsabile a conformarsi agli obblighi internazionali. Una contromisura che sia sproporzionata al danno subito, si rivelerebbe una punizione o ritorsione e quindi sarebbe contraria allo scopo a cui sono destinate le contromisure, con la conseguenza che deve essere considerata illecita, non giustificata.

Gli esperti ritengono peraltro che "proporzionalità" non significhi "reciprocità", nel senso che non vi è alcun requisito che la contromisura di uno Stato danneggiato violi lo stesso obbligo violato dallo Stato responsabile. Né vi è alcun requisito che le contromisure siano della stessa natura dell'atto illecito internazionale sottostante che le giustifica. Contromisure non informatiche possono essere adottate in risposta ad un atto illecito internazionale effettuato con operazioni cyber, e viceversa.

danneggiato”, ossia dallo Stato il cui diritto è stato violato dall’atto illecito internazionale. Le misure adottate da altri Stati, in ausilio allo Stato vittima, possono eventualmente consistere in sanzioni commerciali, finanziarie, e simili. Inoltre, dalla posizione degli esperti internazionali nel Tallinn Manual 2.0, appare pacifico che solo uno Stato può adottare le contromisure, e non le imprese o società private. Ad esempio, un’azienda di tecnologia informatica danneggiata da una operazione cyber che l’ha presa di mira, potrà prendere le proprie precauzioni ed azioni di rimedio contro l’attacco informatico subito, ma la sua risposta non potrà definirsi come una “contromisura”. La contromisura semmai, potrà essere presa dallo Stato di bandiera²³.

Per contro, non vi è alcun divieto per gli Stati danneggiati, di rivolgersi ad un’azienda privata, comprese società straniere, per condurre contromisure informatiche per loro conto, contro gli Stati responsabili. In tal caso, le operazioni della società sarebbero integralmente attribuibili agli Stati danneggiati (secondo i criteri di attribuzioni sopra descritti). Le operazioni informatiche in questione sarebbero pertanto soggette a tutte le restrizioni e condizioni pertinenti sulle contromisure²⁴.

6. Conclusioni e prospettive future

Il campo dell’uso delle ICT da parte degli Stati nel contesto del diritto internazionale è in continua evoluzione. Mentre le tecnologie continuano ad avanzare rapidamente, offrendo enormi vantaggi in termini di comunicazione e gestione dei servizi, il diritto internazionale deve adattarsi ed

²³ Quando la Sony fu oggetto di operazioni cyber dannose nel 2014, presumibilmente condotte dalla Corea del Nord, non aveva il diritto di rispondere all’attacco invocando l’istituto della contromisura in base alla responsabilità dello Stato. Invece il diritto di rispondere con contromisure era riservato agli Stati Uniti (Stato di bandiera della Sony), supponendo che l’operazione contro la Sony potesse essere correttamente definita un atto illecito a livello internazionale da parte della Corea del Nord contro gli Stati Uniti, sulla base della violazione alla sua sovranità. La Sony avrebbe potuto in ogni caso mettere in atto tutte le misure necessarie per bloccare l’attacco nei suoi confronti, ma tali risposte non sarebbero state qualificate come “contromisure”. Qualora Sony avesse risposto all’attacco, non avrebbe violato alcuna regola di diritto consuetudinario internazionale generale, tuttavia, nel farlo, si sarebbe potuto sollevare la violazione del dovere di diligenza da parte degli Stati Uniti. *È stata la Corea del Nord ad attaccare Sony?*, in *ilpost.it*, 20 dicembre 2014

²⁴ Si discute se possano adottare contromisure anche gli Stati non direttamente danneggiati, ma che agiscono a tutela dell’interesse collettivo del gruppo al quale appartengono insieme allo Stato danneggiato; o a tutela del rispetto di obbligazioni generali riconosciute dal diritto consuetudinario (come il divieto di aggressione, di tortura, di genocidio, etc.). Oltre che dall’art. 48 delle ARSIWA e dalle Regole del Tallinn Manual 2.0, rilevante sul punto è la controversia legale internazionale tra Belgio e Spagna, decisa dalla Corte Internazionale di Giustizia (CIJ) nel 1970. Nel caso della *Barcelona Traction*, società costituita in Canada, ma che operava in Spagna, la Corte Internazionale di Giustizia negò che il Belgio (Stato di nazionalità degli azionisti di maggioranza) potesse agire contro la Spagna a tutela della società, affermando che tale diritto spettava allo Stato in cui la società era stata costituita, cioè il Canada - *Caso Barcelona Traction*, deciso dalla Corte Internazionale di Giustizia nel 1970; A. Gentile, *Trattamento delle persone fisiche e giuridiche straniere*, in *ripetiamodiritto.com*,

evolversi per affrontare le nuove sfide e le opportunità che ne conseguono. La protezione della sicurezza informatica, la promozione dei diritti umani online e la prevenzione della criminalità informatica diventano così solo alcune delle aree sulle quali si incentra l'attenzione e che richiedono una continua collaborazione da parte della comunità internazionale.

Per affrontare queste sfide, la comunità internazionale ha sviluppato diverse iniziative multilaterali, cercando di creare un quadro normativo comune volto a delineare le regole per un comportamento responsabile degli Stati nel cyberspazio.

In particolare, l'Organizzazione delle Nazioni Unite, sin dal 2004, con l'istituzione dei gruppi di esperti governativi (Groups of Government Experts, GGE) sulla cybersicurezza, persegue lo scopo di studiare i rischi che si nascondono o potrebbero nascondersi nel cyberspazio, nonché di esaminare le misure collettive che potrebbero essere messe in campo per contrastarli. I sei gruppi di esperti governativi (GGE) che sino ad ora hanno lavorato nel corso degli anni, hanno esaminato le minacce informatiche internazionali ed hanno raggiunto il consenso su alcuni punti fondamentali, che oggi costituiscono i c.d. quattro pilastri su cui si fonda il quadro normativo internazionale del *cyberspace*. In primo luogo, si è confermata l'applicazione del diritto internazionale al comportamento degli Stati nel *cyberspace*, con particolare riferimento all'applicazione della Carta delle Nazioni Unite ed i suoi principi di sovranità, non intervento, risoluzione pacifica delle controversie e rispetto dei diritti umani, principi indispensabili per mantenere la pace e la stabilità internazionale; in secondo luogo, si sono sviluppate delle norme volontarie e non vincolanti (le c.d. "*11 voluntary cyber norms*"), che stabiliscono ciò che gli Stati dovrebbero e non dovrebbero fare nel cyberspazio; in terzo luogo, si sono ideate varie misure volte a rafforzare la fiducia reciproca tra gli Stati, le c.d. "*confidence-building measures*", basate sulla cooperazione, lo scambio di informazioni e la trasparenza delle proprie azioni; ed infine, si sono creati meccanismi di supporto mirati ad aiutare tutti gli Stati ad implementare le capacità informatiche e proteggere i propri networks dagli attacchi informatici, ossia aumentare la propria "*capacity-building*".

Per quanto riguarda le c.d. "*11 voluntary cyber norms*" (poi approvate dall'Assemblea Generale delle Nazioni Unite nel 2015), otto norme fanno riferimento a comportamenti che gli Stati dovrebbero incoraggiare, mentre tre riguardano comportamenti che gli Stati sono chiamati ad evitare. L'obiettivo è quello di favorire innanzitutto la cooperazione tra gli Stati nella crescita di una stabilità nel *cyberspace*, lo scambio di rilevanti informazioni, la prevenzione di incidenti informatici nel proprio territorio, la cooperazione nel contrasto al crimine ed al terrorismo, il rispetto dei diritti umani e della privacy.²⁵

²⁵ Viene infatti stabilito l'obbligo per gli Stati di adottare misure per proteggere le proprie infrastrutture critiche; di collaborare con gli altri Stati per prevenire e mitigare attacchi informatici, condividendo informazioni sulle minacce e rafforzando la resilienza; di incoraggiare pratiche di sicurezza nelle catene di approvvigionamento delle tecnologie per l'informazione e comunicazione (ICT) per prevenire intrusioni dannose; di segnalare eventuali vulnerabilità nei sistemi informatici di altri Paesi, senza sfruttarle per attacchi; di lavorare insieme per costruire capacità nel cyberspazio, migliorando la cooperazione internazionale per la sicurezza informatica. Viene inoltre previsto che gli Stati non

Al momento le c.d. 11 norme per un comportamento responsabile degli Stati, sancite dal GGE nel 2015, hanno carattere “volontario”, ossia non sono giuridicamente vincolanti. In altre parole, gli Stati non hanno l’obbligo giuridico di rispettarle, ma sono incoraggiati ad adottarle su base volontaria, come standard di buona condotta.²⁶

Esse sono peraltro state riprese ed ulteriormente elaborate nei loro aspetti pratici, da un gruppo di lavoro aperto, l’*Open Ended Working Group* (OEWG) delle Nazioni Unite, sulla sicurezza e l’uso delle informazioni e comunicazioni tecnologiche, nel 2021 e, da ultimo, nel report del 2024.²⁷

6.1 La posizione italiana

In risposta alle raccomandazioni rivolte dall’Assemblea Generale delle Nazioni Unite, l’Italia, nel novembre 2021, ha depositato presso l’ONU, il documento relativo alla sua posizione sul diritto internazionale e il cyberspazio (c.d. “*Italian Position Paper on International Law and Cyberspace*”).²⁸ Il documento è stato realizzato dal Ministero degli Affari Esteri, in collaborazione con la Presidenza del Consiglio dei Ministri e il Ministero della Difesa.

Nel suddetto documento l’Italia presenta la sua posizione su molteplici aspetti riguardanti l’applicazione del diritto internazionale nel cyberspazio. Il Paper esordisce facendo proprie le conclusioni raggiunte dal Gruppo di esperti governativi delle Nazioni Unite (GGE) e dal Gruppo di lavoro aperto (OEWG) sulla sicurezza informatica, secondo cui «il diritto internazionale e in particolare la Carta delle Nazioni Unite nella sua interezza, è applicabile ed è essenziale per mantenere la pace e la stabilità e promuovere un ambiente ICT aperto, sicuro, stabile, accessibile e pacifico». L’Italia considera inoltre che, il concetto internazionale di pace e sicurezza vada

dovrebbero consapevolmente permettere che il proprio territorio sia utilizzato per il compimento di atti illeciti internazionali tramite ICT; non dovrebbero condurre attività informatiche contrarie agli obblighi di cui al diritto internazionale che danneggiano le infrastrutture critiche; e non dovrebbero danneggiare i sistemi di risposta alle emergenze informatiche degli altri Paesi (Computer Emergency Team -CERT/ Computer Security Incident Response Team - CSIRT).

²⁶ 2015 UN GGE - *Report of the group of governmental experts on developments in the field of information and telecommunications in the context of international security* ((A/70/174), in *digest*, luglio 2015

²⁷ Questo gruppo è stato istituito dalle Nazioni Unite per promuovere la cooperazione internazionale e per garantire stabilità e sicurezza nel cyberspazio. Sinora ci sono stati due OEWG. Il primo (che ha operato dal 2019 al 2021), ed il secondo (con mandato 2021-2025), che continua sostanzialmente il lavoro del primo OEWG. Ebbene, nel report del 2024, l’OEWG dopo aver dato atto dell’aumento delle campagne malevole per influenzare le politiche degli altri Stati ed in generale dei cyber incidenti, richiama e conferma la validità delle c.d. “*11 voluntary, non-binding norms of responsible State behaviour*”, riconoscendo che ulteriori norme potrebbero essere sviluppate nel tempo. *Open-ended working group on security of and in the use of information and communications technologies*, 2021-2025, Report del 12 luglio 2024 (A/AC.292/2024/CRP.1)

²⁸ F. Zorzi Giustiniani, *Il Position Paper dell’Italia sull’applicabilità del diritto internazionale nel cyberspazio, la sentenza del Tribunale UE nel caso Google Shopping e l’Oxford Statement sulla regolamentazione internazionale degli attacchi ransomware*, in *Nomos*, settembre-dicembre 2021.

oltre la mera dimensione militare, con la conseguenza che le norme di diritto internazionale applicabili nel cyberspazio non sono soltanto quelle relative al divieto dell'uso della forza nelle relazioni internazionali, ma riguardano tutte le attività informatiche. Seppur convinta sull'applicabilità del diritto internazionale al cyberspazio, l'Italia è tuttavia consapevole che adattare le regole e principi esistenti alle caratteristiche del dominio cibernetic, non è agevole. È proprio questo il compito e la sfida che richiede risposte dall'attuale dialogo nella comunità internazionale.

Con riguardo al principio di sovranità, il documento riconosce la sua fondamentale importanza nel cyberspazio, sia nel suo aspetto interno che in quello esterno.

Il documento dedica poi un'intera sezione al diritto della responsabilità degli Stati nel cyberspazio, affrontando la questione dell'attribuzione degli attacchi informatici, il dovere di diligenza (c.d. "*due diligence*"), e la tematica delle contromisure. L'Italia ritiene in primo luogo che l'attribuzione sia una prerogativa nazionale, così come la decisione sul fatto di rendere pubblico o meno il risultato dell'accertamento sull'attribuzione. Nell'affermare la consapevolezza che l'attribuzione comporta considerazioni tecniche, giuridiche e politiche, si ribadisce comunque la necessità di fondarsi su prove sufficientemente affidabili quanto alla fonte delle attività in questione e all'identità dei responsabili. Nel documento si richiamano poi le regole codificate nel "*Progetto di articoli della Commissione del diritto internazionale delle Nazioni Unite del 2001*" (progetto di ARSIWA, esaminato nei capitoli precedenti), riconoscendo alle stesse valore di norme consuetudinarie.

In tema di due diligence, da applicarsi al cyberspazio, il documento richiama quanto affermato dalla Corte Internazionale di Giustizia nel caso di scuola *Corfu Channel* del 1949, nel senso che gli Stati non devono permettere che il loro territorio sia utilizzato per atti contrari al diritto di altri Stati, né tantomeno che la loro infrastruttura ICT (Information and Communication Technology) sia impiegata per la condotta di attività cibernetiche illecite da parte di attori statali o non statali.²⁹ Viene tuttavia ricordato che il dovere di diligenza è un obbligo di condotta e che pertanto uno Stato, fintantoché compie tutti gli sforzi possibili, non può essere ritenuto responsabile se alla fine non è in grado di prevenire, mitigare o porre termine ad attività informatiche illecite avviate o in transito attraverso il suo territorio. Con riguardo alle contromisure, l'Italia riconosce il diritto di ricorrere ad esse, in risposta ad operazioni cibernetiche di altro Stato che costituiscono un illecito internazionale al di sotto della soglia del conflitto armato. Tuttavia, viene dato atto della difficoltà della tracciabilità delle operazioni, dell'individuazione del responsabile degli attacchi e della valutazione della violazione in relazione alla soglia di diligenza dovuta e alla rilevanza del danno subito. Inoltre, adempimenti preventivi quali la previa richiesta di cessazione dell'atto illecito e la previa notifica dell'intenzione di adottare contromisure in risposta alle operazioni informatiche illecite, potrebbero risultare inapplicabili laddove una reazione immediata fosse necessaria per far valere i diritti dello Stato leso e impedire danni ulteriori. In ogni caso le contromisure devono essere proporzionate al danno patito e non possono

²⁹ CIG, *Canale di Corfù (Regno Unito di Gran Bretagna e Irlanda del Nord contro Albania)*, 1949, in *icj-cij.org*.

consistere nella minaccia o nell'uso della forza.

Altra sezione del documento è dedicata alle operazioni informatiche che rientrano nell'ambito del divieto dell'uso della forza ai sensi dell'art. 2 (4) della Carta delle Nazioni Unite. Appare pacifico che, nel caso in cui un'operazione informatica comporti il superamento della soglia (da valutarsi caso per caso) in relazione alla portata e agli effetti in termini di danni fisici alla proprietà, lesioni personali o perdita di vite umane, costituirà "uso della forza". E nel caso in cui l'uso della forza integra gli estremi dell'"attacco armato", vi sarà la possibilità per lo Stato vittima di adottare tutte le risposte in autodifesa, ai sensi dell'art. 51 della Carta delle Nazioni Unite. Viene invece dato atto della sussistenza di perplessità nel caso in cui le operazioni informatiche causano semplicemente la perdita di funzionalità delle infrastrutture. L'inclusione di tali operazioni nel divieto dell'uso della forza, potrebbe giustificarsi se si considera che, a causa dell'importanza dell'affidamento che la società moderna pone sui sistemi informatici e le tecnologie cyber, gli Stati sono da essi dipendenti, con la conseguenza di poter subire gravi danni ed interruzione dei servizi essenziali, anche senza materiali danni fisici.

Altro punto molto importante è il riconoscimento dell'applicabilità nel cyberspazio, del diritto internazionale dei diritti umani. Vi deve essere quindi una sostanziale equiparazione della tutela dei diritti umani sia "on-line" che "off-line", ossia nel mondo virtuale analogamente al mondo reale. In particolare, ogni Stato deve tutelare siffatti diritti, che includono la libertà di opinione e di espressione, il diritto all'accesso alle informazioni e il diritto alla privacy. Al riguardo l'Italia riconosce poi la responsabilità (accountability) anche del settore privato, conformemente ai *Guiding Principles on Business and Human Rights* delle Nazioni Unite del 2011.³⁰

Infine, il documento si chiude con una sezione dedicata all'importanza della cooperazione internazionale, strumento fondamentale nel dominio cibernetico, in ordine al miglioramento della resilienza cibernetica e della stabilità internazionale. Particolare rilievo viene assegnato alle misure di rafforzamento della fiducia tra Paesi, alla condivisione delle informazioni, e alla cooperazione regionale (ad esempio all'interno dell'Unione europea) e bilaterale (ossia con accordi diretti tra due Stati, su questioni specifiche). Considerato inoltre il ruolo di primo piano degli *stakeholders* privati nel cyberspazio, l'Italia considera essenziale la cooperazione pubblico-privato, al fine precipuo di garantire la sicurezza informatica e un efficace rafforzamento delle capacità (capacity-building).

6.2 La posizione europea

All'interno dell'Unione europea, su 27 Stati membri, solo 13 (tra cui, come abbiamo visto, l'Italia) hanno pubblicato la loro posizione ufficiale sul tema dell'applicabilità del diritto internazionale al cyberspazio.

Generalmente, le posizioni nazionali, si basano sul diritto internazionale che trova fondamento nel diritto consuetudinario e nei trattati, come la

³⁰ *Guiding Principles on Business and Human Rights* delle Nazioni Unite, in ohchr.org, 2011

Carte delle Nazioni Unite. Come già specificato, mentre i trattati vincolano solo le parti dello strumento in questione, il diritto consuetudinario stabilisce obblighi legali per tutti gli Stati ed è costituito da regole non scritte ma “cristallizzate” a causa del comportamento o dall’omissione adottata nel corso del tempo con la convinzione di adempiere ad un obbligo legale (*opinio juris ac necessitatis*).

In ogni caso, la singola posizione nazionale e la descrizione di come gli Stati affrontano la tematica del cyberspazio, riveste importanza ed influenza anche nell’interpretazione dei trattati stessi e del diritto consuetudinario. Infatti, maggiore è il consenso su una norma particolare, come ad esempio il rispetto del principio di sovranità e di non intervento negli affari interni ed esterni di altro Stato, più autorevole diventa l’interpretazione di tale norma nel senso percepito dagli Stati stessi.

In ogni caso, lo scorso 18 novembre 2024, il Consiglio dell’Unione europea e i suoi Stati Membri hanno adottato una dichiarazione unanime e comune volta a ribadire che «il diritto internazionale rimane adatto allo scopo in questo settore digitale e [...] gli Stati devono rispettare determinati obblighi e norme nello svolgimento di attività nel ciber spazio».³¹

Con questa dichiarazione l’Unione europea e i suoi Stati membri hanno dimostrato che è possibile raggiungere un’intesa comune su una serie di principi fondamentali e sulle norme di diritto internazionale applicabili al cyberspazio. Una intesa comune globale del modo in cui il diritto internazionale si applica al cyberspazio contribuisce ad aumentare la trasparenza, la prevedibilità e l’assunzione di responsabilità riguardo alla condotta degli Stati nel cyberspazio.

³¹ *Cyberspazio: il Consiglio approva una dichiarazione relativa a un’intesa comune sull’applicazione del diritto internazionale al cyberspazio, in data.consilium.europa.eu, 18 novembre 2024.*

Abstract

In un mondo rivoluzionato dalle nuove tecnologie, gli Stati debbono fare i conti con una nuova realtà, quella cibernetica, e ridefinire i confini internazionali e le frontiere geografiche. Il *cyberspace*, definito come il “*quinto dominio*”, costituisce una realtà virtuale soggetta alla sovranità statale, da disciplinare e tutelare. Ma secondo quali regole? La posizione italiana, depositata nel novembre 2021 presso l’ONU, stabilisce l’applicabilità del diritto internazionale ed in particolare della Carta delle Nazioni Unite nella sua interezza, necessaria per mantenere la pace e la stabilità tra gli Stati.

In a World revolutionized by new technologies, states must grapple with a new reality—cyberspace—and redefine international borders and geographic frontiers. Cyberspace, defined as the “fifth domain”, constitutes a virtual reality subject to state sovereignty, to be regulated and protected. But according to what rules? The Italian position, submitted to the UN in November 2021, establishes the applicability of international law, and in particular the United Nations Charter in its entirety, as necessary to maintain peace and stability among states.

Keywords

uso della forza - cyberspace - diritto internazionale - cybersicurezza - responsabilità internazionale

Elenco autori

Andrea Bertolini

professore associato di diritto privato, Scuola Superiore Sant'Anna

Claudia Calabrese

dottoranda di ricerca in diritto internazionale, Università di Bari "Aldo Moro"

Marco Capece

professore associato di diritto del lavoro, Università di Salerno

Rita Caracuzzo

magistrato

Martina Conticelli

professore ordinario di diritto amministrativo, Università Tor Vergata

Laura Fabiano

professoressa associata di diritto pubblico comparato, Università di Bari "Aldo Moro"

Federica Fedorczyk

post-doctoral research fellow, Oxford University

Angela Ferrari Zumbini

professore ordinario di diritto amministrativo, Università "Federico II" di Napoli

Giacomo Giorgini Pignatiello

culture della materia in Tradizioni costituzionali comparate, Università di Napoli "L'Orientale"

Emma A. Imperato

professoressa associata di diritto pubblico, Università degli Studi di Napoli "L'Orientale"

Marta Infantino

professore associato di diritto privato comparato, Università di Trieste

Carlo Melzi d'Eril

avvocato in Milano

Guilherme Migliora

dottorando di ricerca, Scuola Superiore Sant'Anna

Marta Mariolina Mollicone

assegnista di ricerca in diritto privato, Scuola Superiore Sant'Anna

Federica Paolucci

assegnista di ricerca in diritto costituzionale e pubblico, Università Bocconi

Paolo Passaglia

professore ordinario di diritto pubblico comparato, Università di Pisa

Elisabetta Pino

dottoranda di ricerca in studi sulla criminalità organizzata, Università di Milano

Oreste Pollicino

professore ordinario di diritto costituzionale e pubblico, Università Bocconi

Gabriele Rugani

assegnista di ricerca in diritto dell'Unione europea e diritto pubblico comparato, Università di Pisa

Edoardo Salvatore

dottorando di ricerca in diritto e management della sostenibilità, Università di Siena

Lucia G. Sciannella

professore associato di diritto pubblico comparato, Università di Teramo,

Alice Tadiotto

avvocato in Bologna

Charlotte Waldron

LL.M. graduate in Law of Technology and Automated Systems, Università Bocconi

Nicolò Zammarchi

dottorando di ricerca in diritto penale, Università di Pavia

CODICE ETICO

La **Rivista di diritto dei media** intende garantire la qualità dei contributi scientifici ivi pubblicati. A questo scopo, la direzione, il Comitato degli esperti per la valutazione e gli autori devono agire nel rispetto degli standard internazionali editoriali di carattere etico.

Autori: in sede di invio di un contributo, gli autori sono tenuti a fornire ogni informazione richiesta in base alla policy relativa alle submissions. Fornire informazioni fraudolente o dolosamente false o inesatte costituisce un comportamento contrario a etica. Gli autori garantiscono che i contributi costituiscono interamente opere originali, dando adeguatamente conto dei casi in cui il lavoro o i lavori di terzi sia/siano stati utilizzati. Qualsiasi forma di plagio deve ritenersi inaccettabile. Costituisce parimenti una condotta contraria a etica, oltre che una violazione della policy relativa alle submission, l'invio concomitante dello stesso manoscritto ad altre riviste. Eventuali co-autori devono essere al corrente della submission e approvare la versione finale del contributo prima della sua pubblicazione. Le rassegne di dottrina e giurisprudenza devono dare esaurientemente e accuratamente conto dello stato dell'arte.

Direzione: la direzione (ivi compresi direttori e vice-direttori) si impegna a effettuare la selezione dei contributi esclusivamente in base al relativo valore scientifico. I membri della direzione (ivi compresi direttori e vice-direttori) non potranno fare uso di alcuna delle informazioni acquisite per effetto del loro ruolo in assenza di un'esplicita autorizzazione da parte dell'autore o degli autori. La direzione è tenuta ad attivarsi prontamente nel caso qualsiasi questione etica sia portata alla sua attenzione o emerga in relazione a un contributo inviato per la valutazione ovvero pubblicato.

Comitato degli esperti della valutazione: i contributi sottoposti a valutazione costituiscono documentazione a carattere confidenziale per l'intera durata del processo. Le informazioni o idee acquisite confidenzialmente dai valutatori per effetto del processo di revisione non possono pertanto essere utilizzate per conseguire un vantaggio personale. Le valutazioni devono essere effettuate con profondità di analisi, fornendo commenti e suggerimenti che consentano agli autori di migliorare la qualità delle loro ricerche e dei rispettivi contributi. I revisori dovranno astenersi dal prendere in carico la valutazione di contributi relativi ad argomenti o questioni con i quali sono privi di familiarità e dovranno rispettare la tempistica del processo di valutazione. I revisori dovranno informare la direzione ed evitare di procedere alla valutazione nel caso di conflitto di interessi, derivante per esempio dall'esistenza di perduranti rapporti professionali con l'autore o la relativa istituzione accademica di affiliazione.

