



FIAMMETTA BORGIA\*

## IL MODELLO EUROPEO DI REGOLAZIONE DELL'INTELLIGENZA ARTIFICIALE: TRA AMBIZIONE ETICA E IPERTROFIA NORMATIVA

SOMMARIO: 1. Introduzione – 2. Il quadro internazionale in materia di intelligenza artificiale: limiti dell'approccio generalista e prospettive normative. – 3. La Convenzione quadro del Consiglio d'Europa sull'IA: criticità applicative di un approccio regolatorio etico e orizzontale. – 4. Il modello normativo dell'Unione europea sull'IA: tra idealismo regolatorio e vincoli applicativi. – 5. Conclusioni.

### 1. *Introduzione*

L'intelligenza artificiale (IA) si configura come una tecnologia capace di operare in modo autonomo, assumendo decisioni sulla base dell'analisi dell'ambiente e orientandosi al conseguimento di obiettivi predeterminati<sup>1</sup>. L'evoluzione rapida e pervasiva di tali sistemi, unita alla loro capacità di incidere su ambiti eterogenei e sensibili della vita individuale e collettiva, ha sollevato interrogativi sempre più pressanti circa la necessità di un inquadramento giuridico adeguato. Proprio la natura autonoma, adattiva e, talvolta, opaca dell'IA evidenzia la difficoltà di ricondurre tali tecnologie entro schemi normativi tradizionali, sollecitando gli Stati a interrogarsi su quali strumenti elaborare per affrontare le sfide della cosiddetta “quarta rivoluzione industriale”.

La Comunità internazionale, a fronte della natura transnazionale dell'IA, sta tentando di sviluppare strumenti volti a regolamentare la materia in modo coerente e uniforme, con l'obiettivo di garantire un equilibrio tra innovazione tecnologica e tutela dei diritti fondamentali. Tuttavia, gli sforzi compiuti finora, se da un lato evidenziano una crescente consapevolezza circa la necessità di un'azione normativa globale, dall'altro mettono in luce i limiti intrinseci di un approccio olistico e generalista, finora adottato, rivelando alcune criticità significative.

\* Professoressa ordinaria di diritto internazionale presso il Dipartimento di giurisprudenza dell'Università degli studi di Roma Tor Vergata

<sup>1</sup> Varie sono le definizioni di IA elaborate recentemente. Si veda, ad esempio, S. LEGG-M. HUTTER, *A Collection of Definitions of Intelligence*, in *Frontiers in Artificial Intelligence and Applications*, 2007, 17-24; R.J. STERNBERG, *The Concept of Intelligence*, in R.J. STERNBERG (ed.), *Handbook of Intelligence*, Cambridge, 2012, 3-15.

In primo luogo, si può osservare che il fine di questi strumenti regolatori sia spesso orientato a limitare l'uso dell'IA, piuttosto che a promuoverne la gestione responsabile. Questo approccio difensivo potrebbe avere come effetto non solo quello di rallentare lo sviluppo tecnologico, ma anche impedire di cogliere appieno le opportunità offerte dall'IA, come strumento utile per affrontare problemi globali, come il cambiamento climatico, le pandemie e la povertà. L'enfasi sull'obiettivo della riduzione dei rischi associati all'IA, infatti, ci sembra conduca a un'ipertrofia regolamentare, spesso eccessivamente restrittiva, che può rallentare o addirittura impedire lo sviluppo di tecnologie potenzialmente rivoluzionarie in settori chiave. Inoltre, l'estrema difficoltà di elaborare strumenti vincolanti e a carattere universale in materia, ci sembra impedisca di affrontare efficacemente le sfide transfrontaliere legate all'IA. Infine, l'approccio olistico adottato da molte iniziative regolatorie non sembra tenere conto della complessità e della diversità delle applicazioni dell'IA nei settori più vari, come, ad esempio, la sanità, l'automazione industriale, la finanza o la sicurezza, che necessiterebbero invece di strumenti *ad hoc*. L'approccio olistico e generalista appare, pertanto, inefficace o, peggio ancora, di ostacolo all'innovazione in alcuni settori, senza che sia garantita un'adeguata protezione in altri.

Alla luce di queste considerazioni, l'obiettivo di questo scritto è duplice. In primo luogo, si intende sottolineare la necessità di elaborare un approccio giuridico diverso alla questione, operando lo sforzo di spostare l'obiettivo regolatorio dalla limitazione all'utilizzo responsabile dell'IA. Il fine degli strumenti giuridici internazionali ci sembra, infatti, che debba essere quello di creare un ambiente favorevole per lo sviluppo e l'uso etico dell'IA, piuttosto che quello di limitarne i rischi. Ciò evidentemente richiede un maggiore impegno nella promozione della trasparenza, nella definizione di responsabilità chiare e nella creazione di meccanismi di monitoraggio e cooperazione. In secondo luogo, si vuole sottolineare la necessità di superare l'approccio olistico sino ad ora prevalente per privilegiare una regolamentazione settoriale. Un quadro settoriale ci sembra, infatti, possa essere più idoneo ad affrontare le specificità delle diverse applicazioni dell'IA, garantendo una maggiore effettività normativa.

A tal fine, l'analisi che segue si concentrerà sul cosiddetto "modello europeo" che, pur distinguendosi per l'ambizione etica, appare privo di una visione sufficientemente orientata alla concreta applicazione. La Convenzione quadro del Consiglio d'Europa sull'intelligenza artificiale e il Regolamento dell'Unione europea sull'IA rappresentano, infatti, espressioni emblematiche della marcata ipertrofia regolamentare che caratterizza l'approccio europeo, impegnato in una complessa e faticosa ricerca di equilibrio tra progresso tecnologico e tutela dei diritti umani, più che nella definizione di un quadro normativo funzionale e concretamente applicabile. Proprio l'esame di questi strumenti mostrerà, dunque, i limiti sistematici dell'approccio olistico europeo, sinora valorizzato, evidenziando la necessità di sviluppare piuttosto iniziative regolatorie settoriali, anche superando la "retorica" del rischio.

## *2. Il quadro internazionale in materia di intelligenza artificiale: limiti dell'approccio generalista e prospettive normative*

Nel contesto del diritto internazionale, il principale obiettivo degli Stati in tema di IA sembra essere sino ad ora quello di "limitarne e quindi garantirne" un utilizzo responsabile, in particolare per quanto riguarda quella generativa. Tale tendenza riflette una logica di

contenimento più che di promozione di uno sviluppo normativamente accompagnato e responsabile. Nonostante il carattere frammentario degli interventi in materia, emergono tentativi di elaborare una risposta giuridica sistematica alle sfide poste dall'interazione tra sviluppo tecnologico, strutture democratiche e principi etici

Il dibattito in dottrina, seppur embrionale, è estremamente vitale<sup>2</sup>. Nel tentativo non semplice di individuare alcune direttive giuridiche per il fenomeno in esame, è possibile delineare, seppur con cautela, un quadro normativo strutturato su diversi livelli, che comprende non solo le fonti internazionali in senso stretto, ma anche gli strumenti di *soft law*, i provvedimenti nazionali con impatti transnazionali, le norme tecniche e le misure di autoregolamentazione delle imprese.

In questa prospettiva, andrebbero innanzitutto considerate le numerose norme di diritto internazionale già esistenti che possono essere applicate anche in materia di IA, come quelle a protezione dei diritti umani fondamentali<sup>3</sup>, o relative al diritto internazionale umanitario<sup>4</sup>, nonché alcune convenzioni più recenti, specificamente rivolte a fenomeni legati allo sviluppo tecnologico digitale, ad esempio nell'ambito della protezione dei dati o la Convenzione 108+ del Consiglio d'Europa<sup>5</sup>.

Inoltre, la materia è al centro di numerose iniziative da parte di alcune organizzazioni intergovernative, che si sforzano di configurare un quadro giuridico *ad hoc* per lo spazio digitale<sup>6</sup>. In particolare, dai lavori delle Nazioni Unite emerge l'esigenza di applicare il diritto internazionale alle sfide della digitalizzazione, inclusa l'IA, soprattutto ai fini del mantenimento della pace e della sicurezza internazionale<sup>7</sup>. Tuttavia, tali iniziative, sebbene

<sup>2</sup> Sul punto si veda, tra i più recenti studi, T. BURRI, *International Law and Artificial Intelligence*, in *German Yearbook of International Law*, 2017, 91-108; E. BENVENISTI, *Upholding Democracy Amid the Challenges of New Technology: What Role for the Law of Global Governance?*, in *European Journal of International Law*, 2018, 9-82; M. LILJEFORS-G. NOLL-D. STEUER, *War and Algorithm*, London, 2019; M.M. MAAS, *International Law Does Not Compute: Artificial Intelligence and the Development, Displacement or Destruction of the Global Legal Order*, in *Melbourne Journal of International Law*, 2019, 29-57; P. MOLNAR, *Technology on the Margins: AI and Global Migration Management from a Human Rights Perspective*, in *Cambridge International Law Journal*, 2019, 305-330; A. DEEKS, *Introduction to the Symposium: How Will Artificial Intelligence Affect International Law?*, in *American Journal of International Law Unbound*, 2020, 138-140; M. LANGFORD, *Taming the Digital Leviathan: Automated Decision-Making and International Human Rights*, in *American Journal of International Law Unbound*, 2020, 141-146; M. HILDEBRANDT, *Text-Driven Jurisdiction in Cyberspace*, in *Theoretical and Applied Law*, 2021, 6-20.; A. HÁRS, *AI and International Law: Legal Personality and Avenues for Regulation*, in *Hungarian Journal of Legal Studies*, 2022, 320-344; M. ARVIDSSON-G. NOLL, *Artificial Intelligence, Decision Making and International Law*, in *Nordic Journal of International Law*, 2023, 1-8.

<sup>3</sup> In relazione alle possibili implicazioni dei sistemi di IA sui diritti umani si fa riferimento soprattutto al diritto alla *privacy*, alla libertà di espressione ed al divieto di discriminazione, enunciati – come è noto – nelle disposizioni della Dichiarazione universale dei diritti dell'uomo, della Convenzione europea dei diritti dell'uomo e del Patto sui diritti civili e politici. Si veda sul punto A. BEDUSCHI, *Human Rights and the Governance of Artificial Intelligence*, Geneva Academy of International Humanitarian Law and Human Rights, 2020, reperibile *online*.

<sup>4</sup> H. NASU, *Artificial Intelligence and the Obligation to Respect and to Ensure Respect for IHL*, in E. MASSINGHAM-A. MCCONNACHIE (eds.), *Ensuring Respect for International Humanitarian Law*, I ed., London, 2020.

<sup>5</sup> Al testo originario della Convenzione si affianca il Protocollo n. 223, noto come Convenzione 108+, non ancora entrato in vigore, che comprende nuovi diritti delle persone riguardo a processi decisionali basati su algoritmi, che assumono particolare rilevanza nell'ambito dello sviluppo dell'intelligenza artificiale.

<sup>6</sup> S. FU-A. XING, *International Soft Law Governance of Artificial Intelligence: Advantages, Approaches, and Credibility*, in *Journal of Business Theory and Practice*, 2023, 35-45.

<sup>7</sup> Sul dibattito in seno alle Nazioni Unite sull'intelligenza artificiale v. K.J. SMITH, *The AI Community and the United Nations: A Missing Global Conversation & A Closer Look at Social Good*, in *Association for Advancement of Artificial Intelligence (AAAI) Spring Symposium Series*, 2017, reperibile *online*; E. FOURNIER-TOMBS, *Towards a United Nations Internal Regulation for Artificial Intelligence*, in *Big Data & Society*, 2021, 1 ss.

importanti sul piano dei principi, presentano spesso un contenuto generico e non vincolante, con limitate ricadute operative.

Senza, dunque, avere alcuna pretesa esaustiva, in questa sede ci sembra opportuno almeno richiamare alcune tra le più importanti proposte in tal senso.

Nel sistema delle Nazioni Unite, nel 2021, l'UNESCO ha pubblicato le linee guida sull'uso etico dell'intelligenza artificiale<sup>8</sup>. Queste evidenziano quali principi di riferimento l'equità, la trasparenza, la tutela della *privacy*, la responsabilità e lo sviluppo sostenibile. Nello stesso anno la Conferenza delle Nazioni Unite sul commercio e lo sviluppo (UNCTAD) ha pubblicato una serie di principi sull'intelligenza artificiale e la *governance* dei dati affrontando questioni chiave come il flusso dei dati<sup>9</sup>. Il Consiglio per i diritti umani, nel rapporto del 2021 sull'intelligenza artificiale, ha proposto una serie di principi e linee guida per garantire che le nuove tecnologie basate sull'IA siano sviluppate e applicate in linea con gli *standard* sui diritti umani<sup>10</sup>.

Anche l'Organizzazione per la cooperazione e lo sviluppo economico (OCSE), nel 2019, ha pubblicato una serie di principi di carattere sociale, giuridico ed etico sull'intelligenza artificiale<sup>11</sup>, nei quali si afferma, da un lato, che lo sviluppo e l'applicazione delle nuove tecnologie basate sull'IA dovranno perseguire i principi di giustizia e uguaglianza evitando discriminazioni nei confronti di determinati gruppi o individui, dall'altro, che nei sistemi di IA dovranno essere rispettate le norme relative alla protezione dei dati e dovranno essere adottate misure per garantire la sicurezza e la riservatezza di questi ultimi per prevenire abusi e fughe di dati.

In ambito europeo, il Consiglio d'Europa, nel 2019, su mandato del Comitato dei Ministri, ha istituito un gruppo consultivo *ad hoc* per l'IA al fine di elaborare un accordo in materia di sviluppo, progettazione e utilizzo dell'IA. Nel 2020, il Comitato ha proposto nove principi base per i futuri strumenti giuridici in materia, tra i quali troviamo citati i principi di non-discriminazione, di trasparenza dei sistemi di IA, di salvaguardia della dignità umana, nonché i diritti alla protezione dei dati e alla *privacy*<sup>12</sup>. Più recentemente, nel 2024, sotto l'egida dell'organizzazione è stato concluso il primo trattato sull'IA. La Convenzione quadro del Consiglio d'Europa sull'intelligenza artificiale e i diritti umani, la democrazia e lo Stato di diritto<sup>13</sup> è stata conclusa il 17 maggio 2024 a Strasburgo durante l'incontro ministeriale annuale del Comitato dei Ministri del Consiglio d'Europa. Aperta all'adesione anche da parte di Stati non europei, la Convenzione istituisce un quadro normativo organico applicabile all'intero ciclo di vita dei sistemi di IA, affrontando i potenziali rischi connessi al loro utilizzo e promuovendo, al contempo, un'innovazione tecnologica responsabile. L'accordo adotta

<sup>8</sup> UNESCO, *Recommendation on the Ethics of Artificial Intelligence*, adopted on 23 November 2021, SHS/BIO/PI/2021/1, reperibile *online*.

<sup>9</sup> UNCTAD, *Technology and Innovation Report 2021*, pubblicato il 25 febbraio 2021, UNCTAD/TIR/2020, reperibile *online*.

<sup>10</sup> Human Rights Council, *The right to privacy in the digital age: Report of the United Nations High Commissioner for Human Rights*, 13 September 2021, A/HRC/48/31, reperibile *online*.

<sup>11</sup> OCSE, *Recommendation of the Council on Artificial Intelligence*, adopted on 22 May 2019, OECD/LEGAL/0449 (amended on 03 May 2024).

<sup>12</sup> Ad Hoc Committee On Artificial Intelligence (CAHAI), *Towards Regulation of AI Systems*, December 2020, COE Study DGI (2020)16, reperibile *online*.

<sup>13</sup> Cfr. Committee on Artificial Intelligence (CAI), *Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law*, 17 May 2024, CM(2024)52-final. Tale Convenzione, al momento della pubblicazione di questo scritto, è stata firmata (ma ancora non ratificata) da Andorra, Georgia, Islanda, Montenegro, Norvegia, Moldavia, San Marino, Regno Unito, nonché da Israele, Stati Uniti e dall'Unione europea.

appunto un approccio olistico basato sulla valutazione dei rischi e impone un esame approfondito delle possibili conseguenze negative derivanti dalla progettazione, dallo sviluppo, dall'impiego e dalla dismissione dei sistemi di IA, con il fine di garantire che l'evoluzione tecnologica sia compatibile con la tutela dei diritti fondamentali e con il rispetto dei principi democratici e dello Stato di diritto. Tuttavia, la scelta di un approccio marcatamente olistico e basato sul rischio ne limita fortemente l'efficacia applicativa.

Infine, per quanto riguarda l'Unione europea, nel dicembre 2023, i membri del Consiglio dell'Unione europea e i delegati del Parlamento europeo hanno raggiunto un accordo sulla necessità di creare regole armonizzate sull'IA<sup>14</sup>. Il cosiddetto "Regolamento sull'intelligenza artificiale" ha, infatti, l'obiettivo di assicurare che i sistemi di IA immessi sul mercato europeo e utilizzati nell'UE siano sicuri e rispettino i diritti fondamentali e i valori dell'UE. Si tratta, anche in questo caso, di un'iniziativa legislativa finalizzata a guidare lo sviluppo e l'adozione di un'IA da concepire come "sicura e affidabile", sia da parte di enti pubblici che privati nell'ambito del mercato unico europeo. Il regolamento si basa su un approccio "basato sul rischio", dove la severità delle regole aumenta in relazione al potenziale danno che l'IA può arrecare alla società. Quest'ultima, costituisce – come è noto – la prima iniziativa legislativa di questo genere a livello mondiale e potrebbe stabilire uno *standard* internazionale per la regolamentazione dell'IA in altre regioni, similmente a quanto avvenuto con il Regolamento generale sulla protezione dei dati. Anche in questo caso, però, l'adozione di un modello generalista ed extraterritoriale presenta importanti limiti di effettività.

Quanto agli strumenti di *soft law*, ai provvedimenti nazionali con impatti transnazionali, alle norme tecniche e alle misure di autoregolamentazione delle imprese, che ad oggi completano il quadro internazionale in materia, si tratta soprattutto di strumenti giuridicamente non vincolanti, che concorrono a fornire una guida per la comunità internazionale per promuovere lo sviluppo giuridico ed etico nel campo dell'intelligenza artificiale. Anche in questo caso, gli strumenti richiamati appaiono accomunati da una visione comune che propende per uno sviluppo di un impianto sistematico relativo all'intelligenza artificiale che sia quanto più possibile antropocentrico e riflettono altresì una preoccupazione generalizzata circa i potenziali effetti negativi che l'utilizzo di tali tecnologie potrebbe produrre sulla tutela dei diritti fondamentali<sup>15</sup>. Manca, invece, un'effettiva articolazione operativa che consenta di passare dalla teoria alla prassi.

Da questa sommaria ricognizione dei principali strumenti giuridici esistenti emerge, dunque, come sia a livello universale che nell'ambito di quadri giuridici sovranazionale la normativa in materia di intelligenza artificiale si stia via via delineando secondo un binario ben definito, orientato a limitare più che gestire questa tecnologia (in chiave antropocentrica) e ad offrire medesime (o analoghe) direttive giuridiche per qualsiasi ambito applicativo dell'IA.

La principale debolezza di questo approccio risiede nell'asimmetria tra l'enfasi posta sulla limitazione preventiva e l'elaborazione di strumenti normativi flessibili e applicabili, capaci di adattarsi ai ritmi dell'evoluzione tecnologica. Sul piano giuridico, infatti, l'approccio

<sup>14</sup> Cfr. 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) (Text with EEA relevance), 12.7.2024, EN L series 2024/1689.

<sup>15</sup> Si veda in tal senso C. GRIECO, *Intelligenza artificiale e diritti umani nel diritto internazionale e dell'Unione europea. Alla ricerca di un delicato equilibrio*, in *Ordine internazionale e diritti umani*, 2022, 782-810, in particolare p. 792.

antropocentrico, sebbene fondato su valori fondamentali, tende a cristallizzare una visione difensiva, non considerando adeguatamente la complessità delle interazioni tra IA e sistemi socio-economici globali. Infine, l'idea di fornire soluzioni normative generali per ogni attività svolta grazie all'utilizzo dell'IA risulta poco realistica. Una tale regolamentazione non riesce a cogliere le peculiarità dei diversi settori di applicazione, compromettendo la capacità di rispondere in modo mirato e proporzionato ai rischi concreti posti dalla tecnologia. Questo limita la possibilità di costruire un quadro giuridico efficace e dinamico che possa contemporaneare tutela dei diritti e innovazione tecnologica.

Tali criticità si riflettono chiaramente anche nei più recenti strumenti normativi, come la Convenzione quadro del Consiglio d'Europa sull'intelligenza artificiale e il Regolamento dell'Unione europea sull'intelligenza artificiale. Pur prevedendo una qualche forma di applicazione extraterritoriale – volta, in teoria, a incidere anche al di fuori dei confini strettamente europei – tali strumenti appaiono ancora inadeguati a perseguire concretamente gli ambiziosi obiettivi che si propongono. L'approccio regolatorio si caratterizza infatti per un'eccessiva rigidità e per un'impostazione ipertrofica, poco compatibile con la dinamicità e l'eterogeneità dei contesti globali in cui l'IA si sviluppa. Ne risulta una governance frammentata e poco efficace, incapace di assicurare un reale bilanciamento tra tutela dei diritti fondamentali e promozione dell'innovazione tecnologica su scala globale.

### 3. *La Convenzione quadro del Consiglio d'Europa sull'IA: criticità applicative di un approccio regolatorio etico e orizzontale*

La Convenzione quadro del Consiglio d'Europa sull'IA rappresenta uno strumento normativo di notevole rilevanza per il diritto internazionale contemporaneo, non solo per la sua ambizione di regolamentare un settore in rapida evoluzione, ma anche per il suo tentativo di affrontare, con soluzioni vincolanti per gli Stati contraenti, il delicato equilibrio tra innovazione tecnologica e tutela dei diritti fondamentali.

L'obiettivo, infatti, è quello di regolamentare l'intero ciclo di vita dei sistemi di IA<sup>16</sup>, al fine di garantire la tutela della dignità umana, della democrazia e dello Stato di diritto, come stabilito dall'articolo 3 della Convenzione, e di offrire un bilanciamento tra interessi apparentemente contrapposti. Tuttavia, al di là delle dichiarazioni di intenti, ci sembra – sin da una prima lettura – che si tratti più di una enunciazione di principi che di uno strumento operativo idoneo a fronteggiare le sfide concrete poste dall'IA. Appaiono numerosi, infatti, i dubbi sulla sua effettiva capacità di coniugare progresso tecnologico e tutela dei diritti fondamentali, mentre ci sembra più concreto il rischio di rappresentare un quadro normativo formale che abbia scarsa capacità di essere applicato nella realtà.

Già nel preambolo della Convenzione appare chiaro il suo limite principale: i potenziali benefici delle tecnologie di IA sono appena accennati, mentre i rischi legati al suo utilizzo

<sup>16</sup> Il concetto di “ciclo di vita” (life cycle) dell'intelligenza artificiale include una serie di passaggi, che vanno dalla fase iniziale di preparazione e programmazione, fino alla sua vendita ed eventuale dismissione: «pre-design, design, development, evaluation, testing, deployment, use, sale, procurement, operation and decommissioning». Nel medesimo senso si erano già espresse le Nazioni Unite. Si veda in proposito Risoluzione dell'Assemblea Generale delle Nazioni Unite del 21 marzo 2024, *Seizing the opportunities of safe, secure and trustworthy artificial intelligence systems for sustainable development*, A/RES/78/265, reperibile online.

ispirano l'intera regolamentazione<sup>17</sup>. Infatti, sebbene nel testo siano riconosciute formalmente le opportunità offerte dall'IA per promuovere prosperità e innovazione, le disposizioni successive finiscono per privilegiare una retorica del rischio, imponendo alcuni obblighi stringenti, che possono avere impatti negativi ingiustificati sul suo sviluppo tecnologico<sup>18</sup>. Tale approccio emerge chiaramente nelle misure che vietano l'utilizzo di determinate categorie di IA, quali i sistemi di profilazione, di valutazione sociale, di algoritmi decisionali automatizzati, di riconoscimento biometrico in tempo reale e sistemi predittivi utilizzati per la sicurezza pubblica, ritenute ad alto rischio per i diritti fondamentali e la democrazia, senza prevedere meccanismi adeguati a valutare i benefici che esse potrebbero offrire in altri settori, come la salute, l'istruzione o la sostenibilità ambientale.

Se la visione ispiratrice dello strumento in esame appare sbilanciata, anche l'ambito di applicazione prescelto pone alcuni problemi. La Convenzione, infatti, si riferisce a tutto il ciclo di vita delle intelligenze artificiali che possono influenzare i diritti fondamentali, la democrazia e lo Stato di diritto, rimanendo esclusi i soli sistemi di IA utilizzati per la tutela degli interessi relativi alla sicurezza nazionale, per le attività di ricerca volte a testare i sistemi di IA e per le materie relative alla difesa nazionale<sup>19</sup>. Questa scelta appare contraddittoria rispetto all'ambizione di creare un quadro normativo globale e coerente. Sebbene sia comprensibile la necessità di salvaguardare la sovranità degli Stati membri in settori sensibili, tale esclusione riduce l'efficacia complessiva della Convenzione, lasciando spazio a potenziali lacune normative che potrebbero essere sfruttate dagli Stati stessi o dagli operatori privati per aggirare gli obblighi previsti.

Inoltre, l'approccio adottato dalla Convenzione risente delle difficoltà di armonizzare le esigenze e le sensibilità di quarantasette Stati membri con tradizioni giuridiche differenti. Sebbene sia stato compiuto uno sforzo per stabilire principi universali, come trasparenza,

<sup>17</sup> Si veda in proposito *Council of Europe Framework Convention on Artificial Intelligence and Human Rights*, cit., Preambolo, che dopo aver riconosciuto «that activities within the lifecycle of artificial intelligence systems may offer unprecedented opportunities to protect and promote human rights, democracy and the rule of law», afferma: «Concerned that certain activities within the lifecycle of artificial intelligence systems may undermine human dignity and individual autonomy, human rights, democracy and the rule of law; Concerned about the risks of discrimination in digital contexts [...]»; Concerned by the misuse of artificial intelligence systems and opposing the use of such systems for repressive purposes [...].»

<sup>18</sup> *Ibidem*, art. 8 – Transparency and oversight: «Each Party shall adopt or maintain measures to ensure that adequate transparency and oversight requirements tailored to the specific contexts and risks are in place in respect of activities within the lifecycle of artificial intelligence systems, including with regard to the identification of content generated by artificial intelligence systems»; art. 16 – Risk and impact management framework: «Each Party shall, taking into account the principles set forth in Chapter III, adopt or maintain measures for the identification, assessment, prevention and mitigation of risks posed by artificial intelligence systems by considering actual and potential impacts to human rights, democracy and the rule of law. Such measures shall be graduated and differentiated, as appropriate, and: a) take due account of the context and intended use of artificial intelligence systems, in particular as concerns risks to human rights, democracy, and the rule of law; b) take due account of the severity and probability of potential impacts; c) consider, where appropriate, the perspectives of relevant stakeholders, in particular persons whose rights may be impacted; d) apply iteratively throughout the activities within the lifecycle of the artificial intelligence system; e) include monitoring for risks and adverse impacts to human rights, democracy, and the rule of law; f) include documentation of risks, actual and potential impacts, and the risk management approach; and g) require, where appropriate, testing of artificial intelligence systems before making them available for first use and when they are significantly modified [...].»

<sup>19</sup> *Ibidem*, art. 3, parr. 2, 3, e 4.

responsabilità, sostenibilità e innovazione sicura<sup>20</sup>, la loro realizzazione concreta risulta complessa e frammentata. Ne deriva un quadro normativo inefficace nel rispondere alle sfide globali poste dall'IA, quali, ad esempio, la proliferazione di contenuti manipolati o il rafforzamento delle disuguaglianze sociali. Peraltro, la mancanza di disposizioni specifiche per incentivare lo sviluppo di tecnologie inclusive e sostenibili relegano questi aspetti a dichiarazioni programmatiche, privandoli di immediata efficacia operativa.

Infine, sempre in termini generali, secondo la Convenzione dalla natura intrinsecamente globale dell'intelligenza artificiale deriverebbe la necessità di un approccio multilaterale e collaborativo<sup>21</sup>. Il testo dell'accordo incoraggia gli Stati firmatari non solo a creare normative nazionali coerenti con i principi della Convenzione, ma anche

a promuovere una cooperazione internazionale che faciliti lo scambio di informazioni, esperienze e buone pratiche. La costruzione di una cooperazione tra Stati in materia sarebbe, infatti, fondamentale dal momento che le applicazioni di IA superano i confini nazionali e sollevano problemi di giurisdizione e responsabilità transnazionali. In altri termini, ciò che si vuole proporre è un paradigma di *governance multilivello*, che tenti di coniugare l'autonomia degli Stati con l'esigenza di un approccio coordinato a livello internazionale. Tale impostazione, seppure condivisibile sul piano teorico, presenta criticità evidenti. Basti pensare alla complessità di armonizzare le normative degli Stati firmatari, che produrrà inevitabilmente compromessi applicativi, senza che siano resi effettivi i principi enunciati nella Convenzione. Inoltre, la cooperazione internazionale auspicata, priva di strumenti vincolanti per favorire lo scambio di dati e buone pratiche, ha un'alta probabilità di restare lettera morta, specialmente nell'attuale contesto geopolitico altamente competitivo. Infine, il paradigma di governance multilivello, pur apprezzabile, soffre dell'assenza di soluzioni chiare per affrontare questioni come quelle relative alla giurisdizione e ai regimi di responsabilità. La mancanza di meccanismi concreti di *follow-up* che facilitino la collaborazione tra Stati e l'assenza di sanzioni per la non conformità rende in definitiva il modello proposto più un ideale teorico che un meccanismo operativo efficace.

Quanto al problema del bilanciamento tra innovazione tecnologica e diritti umani, in concreto la Convenzione introduce il concetto di “proporzionalità del rischio”<sup>22</sup>: un principio giuridico che impone di calibrare le misure regolatorie in base alla gravità e alla probabilità dei rischi associati a specifiche applicazioni di IA<sup>23</sup>. Inoltre, è prevista la possibilità, per gli Stati che ne faranno parte, di vietare determinate tipologie di intelligenza artificiale che

<sup>20</sup> *Ibidem*, art. 8 – Transparency and oversight; art. 9 – Accountability and responsibility; art. 12 – Reliability; art. 13 – Safe innovation.

<sup>21</sup> *Ibidem*, art. 25: «The Parties shall co-operate in the realisation of the purpose of this Convention. Parties are further encouraged, as appropriate, to assist States that are not Parties to this Convention in acting consistently with the terms of this Convention and becoming a Party to it [...].».

<sup>22</sup> *Ibidem*, art. 16. Vedi anche C. NOVELLI-F. CASOLARI-A. ROTOLI, et al., *AI Risk Assessment: A Scenario-Based, Proportional Methodology for the AI Act*, in *Digital Society*, 3, 2024. Un approccio analogo è previsto dal Regolamento dell'Unione europea sull'IA.

<sup>23</sup> Tale rischio, tuttavia, a differenza di quanto avviene nell'ambito del diritto dell'Unione europea (dove la gestione del rischio è attuata in via generale ma poi è ulteriormente specificata con la valutazione d'impatto sui diritti fondamentali), si concentra soprattutto sui rischi che l'IA pone per la democrazia, lo Stato di diritto ed i diritti umani [*Council of Europe Framework Convention on Artificial Intelligence and Human Rights*, cit., art. 16, par. 2, lett. a)]. È dunque ragionevole pensare che la gestione del rischio e la valutazione dell'impatto dell'IA su democrazia, Stato di diritto e diritti fondamentali dovrà coordinarsi, per gli Stati membri dell'Unione europea, con quella prevista dall'*AI Act [Regulation (EU) 2024/1689 Of The European Parliament and of the Council of 13 June 2024*, cit., art. 9: Risk management system]. Così in G. ZACCARONI, *Intelligenza artificiale e principio democratico: riflessioni a margine dell'emersione di un quadro normativo europeo*, in *Quaderni AISDUE*, 2, 2024, 1-37, in part. 18.

costituiscono, sulla base degli strumenti di gestione del rischio, un pericolo per la democrazia, i diritti fondamentali e lo Stato di diritto<sup>24</sup>. Questo approccio non solo garantirebbe un'efficace protezione dei diritti fondamentali, ma soffocherebbe l'innovazione tecnologica attraverso regolamentazioni eccessivamente rigide o generiche. Ci sembra, tuttavia che, sebbene il principio sia condivisibile sul piano teorico, la sua applicazione risulti complessa e problematica. Innanzitutto, la valutazione della gravità e della probabilità dei rischi associati alle applicazioni di IA si basa su criteri spesso vaghi e difficili da standardizzare, data la diversità dei contesti tecnologici e sociali in cui queste applicazioni operano. La mancanza di criteri chiari e univoci per identificare e misurare i rischi può portare a interpretazioni soggettive e incoerenti da parte degli Stati contraenti. Inoltre, la calibrazione delle misure regolatorie in base al rischio presuppone un livello elevato di conoscenza tecnica e giuridica da parte delle autorità competenti, il che potrebbe non essere garantito in tutti i contesti nazionali, finendo per generare un'interpretazione e applicazione frammentata e forse non sempre coerente del principio<sup>25</sup>. In definitiva, tale principio rischia di trasformarsi in un'astrazione teorica difficilmente applicabile: l'assenza di un quadro operativo dettagliato e l'elevato margine di discrezionalità lasciato agli Stati rischiano di compromettere seriamente l'efficacia della Convenzione nel garantire il fine del raggiungimento di un bilanciamento tra innovazione tecnologica e tutela dei diritti fondamentali.

Un altro aspetto rilevante della Convenzione è l'enfasi posta sulla trasparenza e sulla intellegibilità degli algoritmi<sup>26</sup>. In un contesto in cui le decisioni automatizzate possono avere un impatto significativo sui diritti individuali, appare, infatti, essenziale che tali decisioni siano comprensibili non solo per gli esperti tecnici, ma anche per gli utenti finali e le autorità di controllo<sup>27</sup>. La Convenzione richiede che gli sviluppatori di IA forniscano una documentazione adeguata e accessibile che descriva il funzionamento dei sistemi, le logiche sottostanti e i potenziali margini di errore. Tale obbligo sarebbe volto a promuovere la responsabilità degli sviluppatori ed a rafforzare la fiducia pubblica nelle tecnologie di IA, un elemento fondamentale per il loro sviluppo e la loro accettazione su larga scala. Anche in tal caso, tuttavia, l'enfasi sulla trasparenza e intellegibilità degli algoritmi, pur fondamentale, presenta limiti significativi. La complessità tecnica di molti modelli di IA, come quelli basati su reti neurali profonde, rende particolarmente complesso offrire una visione pienamente comprensibile delle loro logiche, persino agli esperti. Inoltre, la mancanza di una definizione

<sup>24</sup> Council of Europe Framework Convention on Artificial Intelligence and Human Rights, cit., art. 16, par. 4.

<sup>25</sup> La necessità di bilanciare efficacemente i diritti fondamentali e l'innovazione tecnologica nei settori sensibili, come il sistema giudiziario o la sorveglianza pubblica, appare evidente; è altrettanto chiara la difficoltà di applicare tali principi in questi ambiti. Sebbene l'introduzione di regolamentazioni più stringenti in questi settori possa essere giustificata dalla necessità di tutela dei diritti umani, l'assenza di strumenti concreti per mitigare gli effetti delle restrizioni sullo sviluppo tecnologico potrebbe rischiare di disincentivare gli investimenti in soluzioni innovative, finendo per spingere le imprese tecnologiche a spostarsi verso contesti normativi meno rigidi e riducendo così l'efficacia concreta della Convenzione e la competitività tecnologica degli Stati firmatari.

<sup>26</sup> Council of Europe Framework Convention on Artificial Intelligence and Human Rights, cit., art. 8. I. ERDOGAN, *Diving into the Iceberg: Establishing Transparency in AI for Law Enforcement*, in *European Papers*, 9, 2024, 956-977.

<sup>27</sup> Council of Europe Framework Convention on Artificial Intelligence and Human Rights, cit., art. 14, par. 2, lett. a) e b): «[...] each Party shall adopt or maintain measures including: a) measures to ensure that relevant information regarding artificial intelligence systems which have the potential to significantly affect human rights and their relevant usage is documented, provided to bodies authorized to access that information and, where appropriate and applicable, made available or communicated to affected persons; b) measures to ensure that the information referred to in subparagraph a is sufficient for the affected persons to contest the decision(s) made or substantially informed by the use of the system, and, where relevant and appropriate, the use of the system itself [...].».

chiara dell'espressione "documentazione adeguata", rischia di generare interpretazioni divergenti tra gli Stati, compromettendo l'efficacia delle norme. Infine, l'obbligo di trasparenza impone costi elevati agli sviluppatori, penalizzando in particolare le piccole imprese, e non garantisce automaticamente maggiore fiducia pubblica: la disponibilità di informazioni tecniche non si traduce necessariamente in una loro comprensione da parte di utenti o autorità. Senza *standard* chiari e programmi di sensibilizzazione, la trasparenza rischia di rimanere un principio formale, incapace di promuovere una reale responsabilità.

Appare inoltre degno di nota il tentativo della Convenzione di introdurre il principio dell'innovazione sicura, che consiste nell'utilizzo di ambienti controllati, detti *sandboxes*, per sperimentare e valutare tecnologie di IA in modo circoscritto, prima della loro implementazione su larga scala<sup>28</sup>. Questi ambienti consentono di simulare condizioni reali, identificare potenziali rischi e testare soluzioni in modo sicuro, limitando l'impatto di eventuali problemi sull'ambiente esterno. Al di là dell'obiettivo virtuoso, tuttavia, per garantire l'efficacia di tale approccio, ci sembra sia necessario sviluppare un quadro giuridico e operativo più dettagliato, che includa linee guida pratiche sull'uso delle *sandboxes*, criteri chiari per la valutazione dei risultati e modalità di condivisione delle informazioni ottenute tra gli Stati contraenti. Ad esempio, occorrerebbe definire *standard* uniformi per progettare e condurre i *test*, così da assicurare risultati confrontabili e applicabili in diversi contesti. Inoltre, il principio dell'innovazione sicura dovrebbe essere strettamente collegato a iniziative di cooperazione internazionale. Gli Stati membri potrebbero condividere le conoscenze acquisite nelle *sandboxes*, evitando così duplicazioni di sforzi e frammentazioni normative. Questo approccio collaborativo non solo promuoverebbe uno sviluppo tecnologico più armonioso, ma favorirebbe anche l'adozione di *best practices* comuni, aumentando la fiducia nell'IA e accelerando la realizzazione di soluzioni sicure e innovative a livello globale.

La tutela della privacy e della protezione dei dati personali, riconosciuta come pilastro della Convenzione<sup>29</sup>, si pone in continuità con strumenti come il GDPR definito a livello di UE, introducendo però obblighi specifici per le tecnologie di IA, tra cui le valutazioni d'impatto obbligatorie per i sistemi ad alto rischio<sup>30</sup>. Questo approccio mira a rafforzare l'*accountability* e a garantire che le violazioni dei diritti siano adeguatamente sanzionate. Tuttavia, anche la previsione delle valutazioni d'impatto, sebbene ambiziosa, soffre di vaghezza operativa, proprio legata alla mancanza di una definizione chiara e uniforme di "sistemi ad alto rischio". Un ulteriore limite è l'onere gravoso imposto agli sviluppatori, che potrebbe ostacolare l'innovazione tecnologica, soprattutto per le piccole e medie imprese. Ciò potrebbe incentivare le aziende a spostarsi verso giurisdizioni con regolamentazioni

<sup>28</sup> *Ibidem*, art. 13: «With a view to fostering innovation while avoiding adverse impacts on human rights, democracy and the rule of law, each Party is called upon to enable, as appropriate, the establishment of controlled environments for developing, experimenting and testing artificial intelligence systems under the supervision of its competent authorities». Sulle *sandboxes* (anche se con riferimento all'*AI Act*), vedi G. RUGANI, *La promozione di strumenti di co-regolazione dell'intelligenza artificiale nell'*AI Act*, con particolare riferimento alle regulatory *sandboxes**, in *Quaderni AISDUE*, 2, 2024, 1-15.

<sup>29</sup> *Council of Europe Framework Convention on Artificial Intelligence and Human Rights*, cit., art. 11: « Each Party shall adopt or maintain measures to ensure that, with regard to activities within the lifecycle of artificial intelligence systems: a) privacy rights of individuals and their personal data are protected, including through applicable domestic and international laws, standards and frameworks; and b) effective guarantees and safeguards have been put in place for individuals, in accordance with applicable domestic and international legal obligations».

<sup>30</sup> *Ibidem*, art. 12: «Each Party shall take, as appropriate, measures to promote the reliability of artificial intelligence systems and trust in their outputs, which could include requirements related to adequate quality and security throughout the lifecycle of artificial intelligence systems».

meno rigide, vanificando l'obiettivo di un'applicazione uniforme delle norme in esame<sup>31</sup>. La Convenzione, infine, non affronta adeguatamente il ruolo cruciale della formazione e della sensibilizzazione<sup>32</sup>, elementi indispensabili per una corretta gestione delle tecnologie di IA e per garantire un'effettiva protezione dei dati personali.

La regolamentazione della responsabilità giuridica per i danni causati dai sistemi di IA è, senza dubbio, una delle sfide più complesse per il diritto internazionale. Lo strumento convenzionale in esame dedica uno spazio alla responsabilità giuridica per i danni causati dai sistemi di IA<sup>33</sup>. Sebbene il testo non fornisca una soluzione definitiva, stabilisce principi generali che possono guidare gli Stati nell'elaborazione di normative nazionali coerenti<sup>34</sup>, quali il principio di responsabilità oggettiva per le applicazioni ad alto rischio<sup>35</sup> e l'obbligo di adottare misure preventive per ridurre al minimo i rischi<sup>36</sup>. Questi principi rappresentano un punto di partenza fondamentale per affrontare le complesse questioni di attribuzione della responsabilità in un contesto in cui le decisioni automatizzate possono essere il risultato di interazioni complesse tra diversi attori. La regolamentazione della responsabilità giuridica per i danni causati dai sistemi di IA, pur rappresentando un punto di partenza significativo, presenta diverse criticità. Infatti, come già osservato, la Convenzione si limita a stabilire principi generali, la cui formulazione vaga rischia, anche in tal caso, di generare approcci disomogenei tra gli Stati, creando incertezza giuridica e frammentazione regolatoria. Il principio di responsabilità oggettiva, sebbene utile in astratto, non affronta adeguatamente la complessità delle interazioni tra i vari attori coinvolti nei sistemi di IA, lasciando irrisolta la questione della ripartizione che deriva dall'attribuzione di una responsabilità. Inoltre, l'assenza di meccanismi di armonizzazione tra le normative nazionali potrebbe ostacolare la certezza del diritto e la prevedibilità per gli operatori. Infine, la mancanza di un sistema di *enforcement* efficace e uniforme finisce per compromettere la concreta capacità di garantire la piena attuazione delle misure preventive e di responsabilità<sup>37</sup>. In definitiva, il sistema di responsabilità e di *accountability* delineato appare incompleto, in assenza di strumenti chiari

<sup>31</sup> *Ibidem*, art. 19: «Each Party shall seek to ensure that important questions raised in relation to artificial intelligence systems are, as appropriate, duly considered through public discussion and multistakeholder consultation in the light of social, economic, legal, ethical, environmental and other relevant implications».

<sup>32</sup> *Ibidem*, art. 20: «Each Party shall encourage and promote adequate digital literacy and digital skills for all segments of the population, including specific expert skills for those responsible for the identification, assessment, prevention and mitigation of risks posed by artificial intelligence systems».

<sup>33</sup> *Ibidem*, art. 9: «Each Party shall adopt or maintain measures to ensure accountability and responsibility for adverse impacts on human rights, democracy and the rule of law resulting from activities within the lifecycle of artificial intelligence systems».

<sup>34</sup> *Ibidem*, Chapter III – Principles related to activities within the lifecycle of artificial intelligence systems: artt. 6-13.

<sup>35</sup> *Ibidem*, art. 16 par. 3: «Each Party shall adopt or maintain measures that seek to ensure that adverse impacts of artificial intelligence systems to human rights, democracy, and the rule of law are adequately addressed».

<sup>36</sup> *Ibidem*, art. 16, par. 1.

<sup>37</sup> *Ibidem*, art. 26: «1) Each Party shall establish or designate one or more effective mechanisms to oversee compliance with the obligations in this Convention; 2) Each Party shall ensure that such mechanisms exercise their duties independently and impartially and that they have the necessary powers, expertise and resources to effectively fulfil their tasks of overseeing compliance with the obligations in this Convention, as given effect by the Parties; 3) If a Party has provided for more than one such mechanism, it shall take measures, where practicable, to facilitate effective cooperation among them; 4) If a Party has provided for mechanisms different from existing human rights structures, it shall take measures, where practicable, to promote effective cooperation between the mechanisms referred to in paragraph 1 and those existing domestic human rights structures».

per garantire sanzioni proporzionate e uniformi da parte degli Stati contraenti, aggravato dalla mancanza di meccanismi di monitoraggio efficaci.

I meccanismi di monitoraggio individuati hanno l'obiettivo di valutare l'aderenza degli Stati agli *standard* della Convenzione<sup>38</sup> e includono l'istituzione di una Conferenza delle Parti<sup>39</sup>, che avrà il compito di esaminare periodicamente la realizzazione delle disposizioni convenzionali e di formulare raccomandazioni per il loro miglioramento. Secondo alcuni, la presenza di tali meccanismi può risultare cruciale nel garantire non solo il rispetto formale delle norme, ma anche la loro effettiva applicazione nella pratica<sup>40</sup>. Nella medesima prospettiva, la revisione periodica delle disposizioni della Convenzione consentirebbe di affrontare le sfide poste dalla rapida evoluzione tecnologica, assicurando che il quadro normativo possa rimanere aggiornato e pertinente<sup>41</sup>. Ad una lettura più attenta di queste disposizioni, si può, invece, rilevare che la mancanza di poteri vincolanti riduca l'efficacia delle raccomandazioni della Conferenza delle Parti, lasciando agli Stati un'eccessiva discrezionalità<sup>42</sup>. Inoltre, Inoltre, la Convenzione prevede la possibilità per quest'ultima di proporre al Comitato dei Ministri modalità adeguate al fine di coinvolgere esperti qualificati a supporto di un'attuazione efficace della Convenzione. Sebbene il monitoraggio indipendente sia pensato per garantire una supervisione imparziale, l'imposizione di valutazioni esterne potrebbe essere percepita dagli Stati come un'ingerenza, soprattutto in contesti in cui le normative nazionali divergono significativamente dagli standard della Convenzione<sup>43</sup>. Infine, la revisione periodica delle disposizioni, pur necessaria per rispondere all'evoluzione tecnologica, rischia di risultare inefficace in assenza di procedure chiare per integrare rapidamente le modifiche nei sistemi giuridici nazionali<sup>44</sup>. Senza un coordinamento strutturato tra gli Stati e la Conferenza delle Parti, il processo di aggiornamento regolatorio potrebbe rivelarsi lento o quantomeno parziale, compromettendo la capacità della Convenzione di rimanere al passo con i rapidi sviluppi dell'intelligenza artificiale.

Infine, la Convenzione si coordina con il quadro normativo in evoluzione dell'Unione europea, attribuendole un ruolo significativo: garantisce che l'applicazione della normativa dell'Unione relativa all'intelligenza artificiale avvenga in armonia con la Convenzione, evitando conflitti e riconosce esplicitamente la possibilità per l'Unione di aderire alla

<sup>38</sup> *Ibidem*, Chapter VII – Follow-up mechanism and co-operation: artt. 23-26.

<sup>39</sup> *Ibidem*, art. 23: «The Conference of the Parties shall be composed of representatives of the Parties to this Convention. The Parties shall consult periodically with a view to: a) facilitating the effective application and implementation of this Convention [...]; b) considering the possible supplementation to or amendment of this Convention; c) considering matters and making specific recommendations concerning the interpretation and application of this Convention; d) facilitating the exchange of information on significant legal, policy or technological developments of relevance [...]; e) facilitating, where necessary, the friendly settlement of disputes related to the application of this Convention; and f) facilitating co-operation with relevant stakeholders concerning pertinent aspects of the implementation of this Convention, including through public hearings where appropriate».

<sup>40</sup> A. FARANDA-G. LUSARDI, *Il primo trattato internazionale sull'IA: cosa prevede la convenzione quadro del Consiglio d'Europa*, in *Agenda Digitale*, 28 maggio 2024, reperibile online; F. NIOLA, *l'IA, ecco il primo trattato internazionale: la Convenzione del Consiglio d'Europa*, in *Agenda Digitale*, 6 settembre 2024, reperibile online; M. ODDO, *Intelligenza artificiale, verso regole globali: gli impatti della convenzione quadro*, in *Cybersecurity360*, 12 settembre 2024, reperibile online.

<sup>41</sup> *Council of Europe Framework Convention on Artificial Intelligence and Human Rights*, cit., art. 23, par. 2, lett. a).

<sup>42</sup> *Ibidem*, lett. c).

<sup>43</sup> *Ibidem*, par. 6: «The Conference of the Parties may propose to the Committee of Ministers appropriate ways to engage relevant expertise in support of the effective implementation of this Convention».

<sup>44</sup> *Ibidem*, art. 26.

Convenzione, diventandone formalmente Parte contraente<sup>45</sup>. La Convenzione opera un richiamo al diritto dell'Unione<sup>46</sup>, consentendo di segnalare violazioni che siano al contempo contrarie alla Convenzione e allo stesso diritto unionale, complessivamente inteso<sup>47</sup>. L'espresso riferimento al diritto dell'Unione europea e alla giurisprudenza della Corte di giustizia, peraltro, riflette una volontà di coordinamento nello spazio giuridico europeo che rappresenta un *unicum*.

Anche in questo caso, tuttavia, ci sembra si pongano delle criticità non così facilmente superabili. Infatti, sebbene il riferimento esplicito al diritto dell'Unione e alla giurisprudenza della Corte di giustizia rappresenti un'innovazione, resta ambiguo il modo in cui saranno gestite eventuali discrepanze interpretative, creando potenziali incertezze normative. La possibilità che la Convenzione si configuri come un accordo misto, sottoscritto sia dall'Unione che dagli Stati membri, complica ulteriormente il quadro sul piano delle relazioni tra Unione e suoi Stati membri, in termini di competenze, sia sul piano dello stesso ordinamento UE che sul piano dei rapporti tra UE, Stati membri e Stati terzi europei contraenti la Convenzione, purché la stessa UE non voglia adottare una dichiarazione sulle competenze al momento dell'adesione formale. Inoltre, la Convenzione sembra mancare di un approccio realmente organico alla regolamentazione dell'IA, non affrontando in modo pienamente integrato le molteplici dimensioni coinvolte, come la protezione dei dati, la sicurezza e i diritti fondamentali. Infine, l'integrazione della Convenzione nell'*acquis* dell'Unione potrebbe generare incertezze sulla sua effettiva applicabilità nei procedimenti nazionali.

Dall'analisi che precede si evince come la Convenzione quadro del Consiglio d'Europa sull'intelligenza artificiale rappresenti sicuramente un passo avanti importante verso una regolamentazione internazionale delle tecnologie emergenti, ma sollevi interrogativi significativi sulla sua effettiva efficacia, nonché sulla sua capacità di bilanciare progresso tecnologico e tutela dei diritti fondamentali. Dall'analisi effettuata emergono forti dubbi sulla sua capacità concreta di mitigare i rischi dell'utilizzo dell'IA in assenza di obiettivi mirati, regolamentazione *ad hoc* o settoriale in favore di un approccio olistico che si dimostra utopistico. Inoltre, appare evidente come l'impostazione fortemente precauzionale del testo in esame rischi di limitare lo sviluppo dell'IA, senza offrire garanzie sufficienti per affrontare le sfide globali poste da queste tecnologie.

---

<sup>45</sup> Un passo importante in tal senso è stato compiuto con la firma della Convenzione da parte della Commissione europea il 5 settembre 2024, un gesto che sottolinea non solo la volontà dell'Unione ma anche dei suoi Stati membri di vincolarsi a tale strumento. L'adesione formale dell'Unione renderebbe la Convenzione parte dell'*acquis* comunitario, consentendone l'applicazione anche nei procedimenti giudiziari degli Stati membri. Paradossalmente, l'intera procedura potrebbe concludersi più rapidamente rispetto al percorso, ancora in sospeso, per l'adesione dell'Unione alla Convenzione europea dei diritti dell'uomo. Si veda in proposito G. ZACCARONI, *Intelligenza artificiale e principio democratico*, cit., 11-22.

<sup>46</sup> *Council of Europe Framework Convention on Artificial Intelligence and Human Rights*, cit., art. 27, par. 2: «*Parties which are members of the European Union shall, in their mutual relations, apply European Union rules governing the matters within the scope of this Convention without prejudice to the object and purpose of this Convention and without prejudice to its full application with other Parties. The same applies to other Parties to the extent that they are bound by such rules.*»

<sup>47</sup> Cfr. *Explanatory Report to the Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law*, 5 September 2024, CETS 225, par. 147: «*[...] measures, principles and procedures provided for in the European Union legal order, in particular laws, regulations or administrative provisions as well as other requirements, including court decisions [...]».*

#### 4. Il modello normativo dell'Unione europea sull'IA: tra idealismo regolatorio e vincoli applicativi

La regolamentazione dell'Unione Europea in materia di IA si colloca all'interno di una strategia ampia e ambiziosa volta a bilanciare l'innovazione tecnologica con la tutela dei diritti fondamentali, della sicurezza e della trasparenza. Come è noto, al centro di questa iniziativa normativa vi è il Regolamento sull'Intelligenza Artificiale (*AI Act*)<sup>48</sup>, proposto dalla Commissione Europea nell'aprile 2021, che rappresenta il primo tentativo organico di disciplinare l'IA<sup>49</sup>. Oltre all'*AI Act*, la materia è integrata da altre iniziative legislative europee, come il Regolamento generale sulla protezione dei dati (GDPR)<sup>50</sup>, che stabilisce standard elevati per il trattamento dei dati personali utilizzati dalle tecnologie di IA, il *Digital Services Act* (DSA)<sup>51</sup> e il *Digital Markets Act* (DMA)<sup>52</sup>, che affrontano le responsabilità delle piattaforme digitali e delle *big tech*, promuovendo un ambiente digitale sicuro e competitivo<sup>53</sup>. Da ultimo, si deve segnalare che è in corso di preparazione una direttiva relativa all'adeguamento delle norme in materia di responsabilità civile extracontrattuale all'IA<sup>54</sup>. La Commissione Europea ha anche lanciato il Piano Coordinato sull'IA<sup>55</sup>, che mira a promuovere la cooperazione tra Stati membri per lo sviluppo e l'implementazione dell'IA, e il programma *Horizon Europe*<sup>56</sup>, che finanzia la ricerca e l'innovazione nel campo dell'intelligenza artificiale.

L'analisi che segue non ha come obiettivo di focalizzarsi sulle specifiche soluzioni giuridiche elaborate dagli strumenti richiamati, oggetto di approfondite riflessioni in altri contributi di questo *Focus*, ma di evidenziarne gli aspetti di criticità che emergono, sempre

<sup>48</sup> Regulation (EU) 2024/1689 Of The European Parliament and of the Council of 13 June 2024, cit.

<sup>49</sup> Il regolamento sull'intelligenza artificiale (*AI Act*) è stato estesamente commentato. Si vedano da ultimo i contributi inseriti nei *Quaderni AISDUE* - Rivista quadrimestrale ISSN 2975-2698 - Fascicolo speciale n. 2/2024 La nuova disciplina UE sull'intelligenza artificiale, e in particolare: E. CIRONE, *L'AI Act e l'obiettivo (mancato?) di promuovere uno standard globale per la tutela dei diritti fondamentali*; F. FERRI, *Il giorno dopo la rivoluzione: prospettive di attuazione del regolamento sull'intelligenza artificiale e poteri della Commissione Europea*; M. INGLESE, *Il regolamento sull'intelligenza artificiale come atto per il completamento e il buon funzionamento del mercato interno?*; S. VILLANI, *Il sistema di vigilanza sull'applicazione dell'AI Act: ognun per sé?*; nonché C. NOVELLI, *L'Artificial Intelligence Act europeo: alcune questioni di implementazione*, in *federalismi.it*, 2024, 95 ss.; C. NOVELLI-F. CASOLARI-A. ROTOLI-M. TADDEO-L. FLORIDI, *Taking AI risks seriously: a new assessment model for the AI Act*, in *AI & Society*, 2023, 1 ss.; J. SCHUETT, *Risk Management in the Artificial Intelligence Act*, in *European Journal of Risk Regulation*, 2023, 1 ss.; M. ALMADA-N. PETIT, *The EU AI Act: a Medley of Product Safety and Fundamental Rights?*, in *Robert Schuman Centre Working Paper*, 2023, reperibile online.

<sup>50</sup> Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons 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), OJ L 119, 4.5.2016, Corrigendum, OJ L 127, 23.5.2018.

<sup>51</sup> 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), L 277/1, 27.10.2022.

<sup>52</sup> Regulation (EU) 2022/1925 of the European Parliament and of the Council of 14 September 2022 on contestable and fair markets in the digital sector and amending Directives (EU) 2019/1937 and (EU) 2020/1828 (Digital Markets Act), L 265/1, 12.10.2022.

<sup>53</sup> Per un'analisi sistematica degli strumenti in esame si veda C. SCHEPISI, *Le "dimensioni" della regolazione dell'intelligenza artificiale nella proposta di regolamento della Commissione*, in *Quaderni AISDUE, Sezione "Atti convegni AISDUE"*, 16, 2022, 347 ss.

<sup>54</sup> *Proposal for a Directive of the European Parliament and of the Council on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive)*, COM/2022/496 final, reperibile online.

<sup>55</sup> Commissione Europea, *Piano coordinato sull'intelligenza artificiale*, COM/2018/795 final, reperibile online.

<sup>56</sup> European Commission: Directorate-General for Research and Innovation, *Horizon Europe, open science – Early knowledge and data sharing, and open collaboration*, Publications Office of the European Union, 2021, reperibile online.

con il fine di mostrare come anche il modello unionale, per quanto pionieristico, risenta di impostazioni logico-giuridiche a tratti utopistiche e si scontri nella sua applicazione con ostacoli sistematici difficilmente superabili<sup>57</sup>.

In tale prospettiva, dopo aver effettuato alcune considerazioni preliminari in merito alla finalità, all'approccio e alla definizione stessa di intelligenza artificiale contenuta nell'*AI Act*, ci si concentrerà sui due aspetti chiave dello strumento in esame: la questione della tutela dei diritti umani e l'approccio basato sul rischio.

L'esame complessivo del Regolamento mostra innanzitutto l'intenzione del legislatore dell'Unione europea di creare una regolamentazione con un raggio d'azione che si estenda oltre i confini unionali e, in secondo luogo, la volontà di adottare volutamente un approccio orizzontale, che non si concentri su problematiche settoriali o specifiche, ma miri a disciplinare trasversalmente tutti i settori interessati dall'intelligenza artificiale. A ciò si aggiunga l'incerta formulazione della definizione di IA volta a ricoprendere un ampio spettro di usi della stessa.

Sia l'obiettivo perseguito che l'approccio adottato risultano discutibili. Quanto all'ambizione di progettare il quadro normativo oltre i confini dell'Unione, pur rappresentando un tentativo di acquisire una posizione di "supremazia normativa" a livello globale, rischia di rivelarsi problematica. La regolamentazione UE, ispirata da valori come la protezione dei diritti fondamentali e la sicurezza, potrebbe entrare in conflitto con le priorità normative e commerciali di altre giurisdizioni, come quelle di Stati Uniti, Cina e India, che adottano approcci più flessibili o orientati alla promozione dell'innovazione, creando tensioni economiche e politiche e ostacolando la cooperazione globale in questo ambito. Inoltre, l'imposizione di standard UE a operatori esterni potrebbe indurre alcune imprese a evitare il mercato europeo, con conseguenti ripercussioni negative sulla competitività dell'industria tecnologica nell'UE.

Con riferimento all'adozione di un approccio orizzontale, volto a disciplinare in modo uniforme tutti i settori interessati dall'intelligenza artificiale, quest'ultimo pur garantendo coerenza normativa, si scontra con la complessità e la diversità delle applicazioni dell'IA. Questo approccio, ispirato al modello del GDPR, mira a evitare frammentazioni normative tra gli Stati membri e a garantire che ogni utilizzo dell'IA rispetti standard comuni di tutela dei diritti fondamentali. Tuttavia, questa impostazione solleva interrogativi circa la sua efficacia pratica. La pluralità degli ambiti di applicazione dell'IA rende difficile stabilire requisiti univoci che siano al contempo adeguati e sufficientemente specifici<sup>58</sup>. Inoltre, la regolamentazione orizzontale rischia di creare sovrapposizioni con altre normative settoriali, generando ulteriore complessità non solo per gli operatori giuridici, ma soprattutto per gli operatori del mercato. Infine, l'integrazione di normative IA con i regolamenti già esistenti (ad esempio nel settore sanitario o dei trasporti) richiederà un coordinamento significativo tra le autorità nazionali e sovranazionali, con il rischio di rallentare l'attuazione delle nuove regole.

Quanto al concetto stesso di intelligenza artificiale, quest'ultima è concepita in chiave metagiuridica come «una famiglia di tecnologie in rapida evoluzione che contribuisce al conseguimento di un'ampia gamma di benefici a livello economico, ambientale e sociale

<sup>57</sup> C. NOVELLI, *L'Artificial Intelligence Act Europeo*, cit., 102-105.

<sup>58</sup> Ad esempio, l'art. 13 dell'*AI Act* stabilisce obblighi generali di trasparenza per tutti i sistemi di IA, ma non tiene conto delle differenze operative tra una piattaforma di *e-commerce* e un sistema di sorveglianza pubblico: «High-risk AI systems shall be designed and developed in such a way as to ensure that their operation is sufficiently transparent to enable deployers to interpret a system's output and use it appropriately [...].».

nell'intero spettro delle attività industriali e sociali<sup>59</sup>. Si tratta di una definizione, ampia e generica, che riflette la natura fluida del concetto stesso di IA e l'intento di adattarsi ai rapidi sviluppi tecnologici e alla varietà delle sue applicazioni. Tuttavia, tale scelta definitoria sebbene abbia il pregio di riconoscere la natura eterogenea e dinamica dell'IA, solleva inevitabilmente difficoltà nella traduzione pratica dei principi normativi. In secondo luogo, questa impostazione metagiuridica sembra confermare l'intenzione dell'Unione Europea di creare un quadro normativo universale, capace di adattarsi all'evoluzione costante della tecnologia. Sebbene questa flessibilità possa sembrare vantaggiosa per affrontare il carattere dinamico dell'IA, essa comporta il rischio di un'applicazione normativa frammentata o incoerente nella prassi degli Stati membri. Inoltre, la scelta di una definizione così inclusiva, pur ponendosi l'obiettivo di coprire un ampio spettro di applicazioni, tende a diluire l'attenzione su questioni specifiche e critiche, come la trasparenza degli algoritmi, la protezione dei dati personali, la responsabilità giuridica per danni causati dall'IA e la prevenzione di discriminazioni. Infine, l'adozione di una definizione così generale riflette anche la complessità di elaborare una normativa che possa essere applicabile non solo all'interno dell'Unione, ma anche a livello globale.

Quanto ai principali aspetti del Regolamento, la preoccupazione per la possibilità di compressione dei diritti fondamentali e l'approccio basato sul rischio ne costituiscono i pilastri fondamentali.

Innanzitutto, come la Convenzione del Consiglio d'Europa, anche lo strumento in esame pone l'accento sulla protezione dei diritti fondamentali, considerandola l'obiettivo principale della normativa<sup>60</sup>. Questo principio emerge chiaramente nell'articolato del Regolamento, dove l'importanza della salvaguardia della dignità umana, della *privacy* e della non discriminazione viene ribadita più volte, in linea con gli *standard* già stabiliti in altri strumenti del diritto dell'Unione Europea<sup>61</sup>, come la Carta dei Diritti Fondamentali dell'UE<sup>62</sup> e lo stesso Regolamento Generale sulla Protezione dei Dati (GDPR)<sup>63</sup>. Ampio spazio peraltro è riservato al problema della trasparenza. Quest'ultima è intesa come la capacità di garantire che i sistemi di IA siano progettati e utilizzati in modo tracciabile e spiegabile, informando adeguatamente sia gli utenti finali che i responsabili del loro impiego, inclusa la

<sup>59</sup> *Ibidem*, Considerando 4: «AI is a fast evolving family of technologies that contributes to a wide array of economic, environmental and societal benefits across the entire spectrum of industries and social activities [...].»

<sup>60</sup> A. ADINOLFI, *L'unione europea dinanzi allo sviluppo dell'intelligenza artificiale: la costruzione di uno schema di regolamentazione europeo tra mercato unico digitale e tutela dei diritti fondamentali*, in S. DORIGO (a cura di), *Il ragionamento giuridico nell'era della intelligenza artificiale*, Pisa, 2020, 14 ss.

<sup>61</sup> Si vedano sul punto le riflessioni di G. CONTALDI, *La proposta di regolamento sull'intelligenza artificiale e la protezione dei dati personali*, in G. CAGGIANO-G. CONTALDI-P. MANZINI (a cura di), *Verso una legislazione europea su mercati e servizi digitali*, Bari, 2021, 207 ss.; ma anche C. NOVELLI-F. CASOLARI-P. HACKER-G. SPEDICATO-L. FLORIDI, *Generative AI in EU Law: Liability, Privacy, Intellectual Property, and Cybersecurity*, in *Computer Law and Security Review*, 55, 2024, 1-16.

<sup>62</sup> Gazzetta ufficiale dell'Unione europea, *Carta dei diritti fondamentali dell'Unione europea*, 2016/C 202/02, 389-405. L'utilizzo dell'IA con le sue caratteristiche specifiche (ad esempio opacità, complessità, dipendenza dai dati, comportamento autonomo) può incidere negativamente su una serie di diritti fondamentali sanciti dalla Carta dei diritti fondamentali dell'Unione europea. Si veda in proposito Commissione Europea, *Proposta di regolamento del Parlamento europeo e del Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale (legge sull'intelligenza artificiale) e modifica alcuni atti legislativi dell'Unione* – Relazione, COM/2021/206 final, punto 3.5; Commissione Europea, *Creare fiducia nell'intelligenza artificiale antropocentrica*, COM/2019/168 final; *Risoluzione del Parlamento europeo del 20 ottobre 2020 recante raccomandazioni alla Commissione concernenti il quadro relativo agli aspetti etici dell'intelligenza artificiale, della robotica e delle tecnologie correlate*, 2020/2012(INL).

<sup>63</sup> *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016*, cit., art. 1.

comunicazione sui limiti e le potenzialità del sistema, oltre ai diritti delle persone coinvolte. Il Regolamento, infine, amplia il discorso oltre la protezione dei diritti fondamentali, assumendo una prospettiva globale che collega l'intelligenza artificiale a pratiche commerciali scorrette, come definite nella direttiva 2005/29/CE<sup>64</sup>. Nel Considerando 29<sup>65</sup>, si evidenzia come l'IA possa essere utilizzata per manipolare le persone, inducendole ad adottare comportamenti indesiderati o a prendere decisioni in modo ingannevole, compromettendo la loro autonomia e capacità decisionale.

Uno degli elementi distintivi dell'*AI Act* è l'adozione di misure specificamente progettate per mitigare i rischi che l'intelligenza artificiale può rappresentare per i diritti fondamentali<sup>66</sup>. In questo senso, un aspetto fondamentale della normativa è l'adozione del cosiddetto approccio basato sul rischio<sup>67</sup>, simile a quello utilizzato in altre normative europee, come il GDPR<sup>68</sup> o gli atti normativi in materia di antiriciclaggio<sup>69</sup>. Tale approccio è utilizzato per classificare i sistemi di IA in base al loro potenziale di arrecare danno ai diritti fondamentali, sottponendo ciascuno di essi a regimi normativi differenti. L'art. 6 distingue i sistemi ad alto rischio, che includono applicazioni in settori come la sanità, la giustizia e il lavoro, e quelli a rischio inaccettabile<sup>70</sup>, come i sistemi di manipolazione subliminale o di sorveglianza biometrica in tempo reale, vietati dall'art. 5<sup>71</sup>. Alcuni sistemi di IA sono vietati all'interno dell'Unione perché considerati lesivi della dignità umana o di diritti fondamentali, mentre altri sono classificati come ad alto rischio sulla base di criteri specifici. Tra questi, si annoverano: lo scopo del sistema di IA, la quantità e la natura dei dati trattati, il grado di autonomia del sistema, la possibilità di intervento umano per mitigare decisioni potenzialmente dannose, la probabilità e l'entità dei danni causati e la vulnerabilità delle persone coinvolte, in relazione a fattori come età, condizione sociale o economica, o squilibri di potere rispetto agli utilizzatori del sistema.

<sup>64</sup> Direttiva 2005/29/CE del Parlamento europeo e del Consiglio dell'11 maggio 2005 relativa alle pratiche commerciali sleali tra imprese e consumatori nel mercato interno e che modifica la direttiva 84/450/CEE del Consiglio e le direttive 97/7/CE, 98/27/CE e 2002/65/CE del Parlamento europeo e del Consiglio e il regolamento (CE) n. 2006/2004 del Parlamento europeo e del Consiglio («direttiva sulle pratiche commerciali sleali»), L 149/22, 11.6.2005.

<sup>65</sup> Regulation (EU) 2024/1689 Of The European Parliament and of the Council of 13 June 2024, cit., Considerando 29: «AI-enabled manipulative techniques can be used to persuade persons to engage in unwanted behaviours, or to deceive them by nudging them into decisions in a way that subverts and impairs their autonomy, decision-making and free choices».

<sup>66</sup> C. SCHEPISI, *Diritti fondamentali, principi democratici e rule of law: quale ruolo e quale responsabilità per gli Stati nella regolazione dell'intelligenza artificiale*, in *Studi sull'integrazione europea*, 2022, 41-66.

<sup>67</sup> Su tale sistema, e sulle differenze e gradazioni nei diversi atti legislativi dell'Unione che lo applicano, vedi G. DE GREGORIO-P. DUNN, *The European Risk-Based Approaches: Connecting Constitutional Dots in the Digital Age*, in *Common Market Law Review*, 2022, 473-500.

<sup>68</sup> Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016, cit., art. 35: Data protection impact assessment.

<sup>69</sup> Regolamento (UE) 2024/1620 del Parlamento europeo e del Consiglio del 31 maggio 2024 che istituisce l'Autorità per la lotta al riciclaggio e al finanziamento del terrorismo e che modifica i regolamenti (UE) n. 1093/2010, (UE) n. 1094/2010 e (UE) n. 1095/2010, GU L del 19.6.2024, art. 8: Metodologia di supervisione AML/CFT.

<sup>70</sup> Regulation (EU) 2024/1689 Of The European Parliament and of the Council of 13 June 2024, cit., art. 6: «Irrespective of whether an AI system is placed on the market or put into service independently of the products referred to in points (a) and (b), that AI system shall be considered to be high-risk where both of the following conditions are fulfilled: (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 harmonization legislation listed in Annex I; (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 harmonization legislation listed in Annex I».

<sup>71</sup> *Ibidem*, art. 5, par. 1, lett. a), b).

Sebbene questa impostazione consenta una regolamentazione flessibile, essa pone alcuni dubbi, soprattutto di ordine applicativo. Innanzitutto, la definizione stessa di “rischio” è problematica. L’ampiezza della definizione dei sistemi di IA, che include tutte le tecnologie basate su apprendimento automatico, ragionamento logico e statistico, rischia di generare incertezze interpretative. Questo potrebbe portare a discrepanze nell’applicazione del regolamento tra gli Stati membri, vanificando l’obiettivo di un’armonizzazione normativa. Inoltre, il Regolamento non chiarisce sufficientemente i criteri per determinare il livello di rischio. L’art. 14, ad esempio, impone che i sistemi ad alto rischio siano soggetti a supervisione umana per mitigare i rischi imprevisti, ma non specifica quali siano gli standard per valutare tali rischi<sup>72</sup>. Questa lacuna evidenzia un paradosso intrinseco: mentre l’approccio precauzionale mira a prevenire i danni, l’incertezza normativa potrebbe frenare l’innovazione, scoraggiando le imprese dall’investire in tecnologie IA all’interno dell’Unione Europea. Infine, il Regolamento prevede che i sistemi di IA che non rientrano nelle categorie ad alto rischio non siano soggetti a obblighi stringenti, a meno che non presentino rischi significativi per la salute, la sicurezza o i diritti fondamentali, o non influenzino in maniera rilevante l’esito di processi decisionali. Questo approccio graduato mira a bilanciare l’innovazione tecnologica con la protezione degli individui, evitando un’applicazione normativa eccessivamente onerosa per i sistemi a rischio limitato.

Il Regolamento introduce, inoltre, la valutazione di impatto sui diritti fondamentali come elemento chiave per garantire che l’introduzione e l’utilizzo dei sistemi di IA siano conformi agli *standard* etici e legali europei. Ai sensi del primo paragrafo dell’art. 27, la valutazione di impatto è articolata in sei sezioni principali: a) una descrizione dei processi nei quali il sistema verrà impiegato, b) una definizione del tempo e della frequenza di utilizzo, c) l’individuazione delle categorie di persone fisiche e gruppi che saranno potenzialmente interessati nel contesto specifico, d) l’analisi dei rischi specifici per tali categorie, e) una spiegazione delle modalità di sorveglianza umana previste, e f) le misure progettate per mitigare i rischi identificati<sup>73</sup>.

Tuttavia, anche questo sistema presenta alcune criticità che meritano attenzione. Innanzitutto, non è chiaro il rapporto tra questa valutazione di impatto e il più ampio sistema di gestione del rischio previsto dal Regolamento. Sebbene il secondo paragrafo dell’art. 27 sembri indicare che si tratti di due processi distinti<sup>74</sup>, esiste il rischio di sovrapposizioni o di confusione tra i vari obblighi, specialmente quando il *deployer* deve integrare valutazioni già effettuate dal fornitore o previste da altre normative. Questa complessità potrebbe aumentare gli oneri amministrativi per gli operatori, soprattutto per quelli meno strutturati, come le

<sup>72</sup> *Ibidem*, art. 14: «High-risk AI systems shall be designed and developed in such a way, including with appropriate human-machine interface tools, that they can be effectively overseen by natural persons during the period in which they are in use. Human oversight shall aim to prevent or minimize the risks to health, safety or fundamental rights that may emerge when a high-risk AI system is used in accordance with its intended purpose or under conditions of reasonably foreseeable misuse [...].»

<sup>73</sup> *Ibidem*, art. 27, par. 1: «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 [...].»

<sup>74</sup> *Ibidem*, par. 2: «The obligation laid down in paragraph 1 applies to the first use of the high-risk AI system. The deployer may, in similar cases, rely on previously conducted fundamental rights impact assessments or existing impact assessments carried out by provider. If, during the use of the high-risk AI system, the deployer considers that any of the elements listed in paragraph 1 has changed or is no longer up to date, the deployer shall take the necessary steps to update the information».»

piccole e medie imprese. Un'ulteriore criticità riguarda l'effettiva qualità e uniformità delle valutazioni di impatto. L'assenza di linee guida dettagliate su come condurre queste valutazioni potrebbe portare a interpretazioni divergenti tra i diversi operatori o Stati membri, compromettendo l'obiettivo di coerenza e armonizzazione del Regolamento. Inoltre, il rischio che le valutazioni siano trattate come meri adempimenti burocratici, piuttosto che come strumenti per identificare e mitigare rischi reali, potrebbe limitarne l'efficacia pratica. Un altro elemento problematico è legato alla notifica dei risultati all'autorità di vigilanza del mercato. Sebbene questa misura miri a rafforzare la trasparenza e la supervisione, non è chiaro se le autorità disporranno delle risorse e delle competenze necessarie per analizzare e verificare un numero potenzialmente elevato di valutazioni. Ciò potrebbe creare ritardi o disfunzioni operative, riducendo l'impatto concreto della regolamentazione. Infine, il quinto paragrafo dell'art. 27 assegna all'Ufficio per l'Intelligenza Artificiale il compito di predisporre un questionario semplificato per agevolare i *deployer*<sup>75</sup>, sebbene resti incerta l'effettiva utilità di questo strumento, poiché la semplificazione potrebbe compromettere la capacità di affrontare rischi complessi.

Il Regolamento delinea in definitiva un quadro giuridico complesso, centrato sulla protezione dei diritti fondamentali attraverso la gestione del rischio e la valutazione dell'impatto. Ci sembra, tuttavia, che sia necessario operare qualche riflessione, sia sul piano concettuale che pratico.

In primo luogo, come già rilevato l'approccio adottato risulta eccessivamente generalista. La scelta di disciplinare l'intero ecosistema dell'intelligenza artificiale con un unico quadro regolatorio, senza considerare le peculiarità dei vari settori, rischia di trasformare il Regolamento in una raccolta di enunciazioni di principio, più che in un insieme di norme chiare, precise ed immediatamente applicabili senza ulteriori precisazioni. La varietà delle applicazioni dell'IA richiederebbe una regolamentazione più specifica e settoriale, in grado di rispondere alle sfide proprie di ciascun contesto tecnologico e sociale. La mancanza di tale differenziazione non solo complica l'applicazione pratica delle norme, ma aumenta il rischio di interpretazioni divergenti tra gli Stati membri, compromettendo l'armonizzazione normativa all'interno dell'Unione.

Infine, l'approccio precauzionale del Regolamento, che impone rigidi obblighi di conformità e complessi meccanismi di valutazione del rischio, rischia di soffocare lo sviluppo tecnologico delle aziende europee. Sebbene l'intento dell'Unione sia quello di estendere l'efficacia normativa del Regolamento oltre i confini europei, con una portata extraterritoriale, nella pratica questa ambizione potrebbe rivelarsi controproducente. Le imprese europee, costrette a operare in un contesto normativo stringente e burocratizzato, si troverebbero in una posizione di svantaggio competitivo rispetto a quelle che operano in giurisdizioni con regole più flessibili, come Stati Uniti, Cina o India. Questo scenario rischia di penalizzare l'innovazione e gli investimenti nel settore tecnologico europeo, incentivando potenzialmente una delocalizzazione delle attività verso contesti regolatori meno restrittivi.

In definitiva, il Regolamento, pur avendo l'obiettivo di garantire la tutela dei diritti fondamentali e la sicurezza nell'uso dell'intelligenza artificiale, rischia di fallire nel bilanciare adeguatamente queste priorità con la necessità di promuovere lo sviluppo tecnologico e la competitività dell'industria europea. Anche in UE, ci sembra che un approccio più mirato,

<sup>75</sup> *Ibidem*, par. 5: «The AI Office shall develop a template for a questionnaire, including through an automated tool, to facilitate deployers in complying with their obligations under this Article in a simplified manner».

flessibile e calibrato sui diversi contesti applicativi potrebbe offrire una soluzione più efficace per affrontare le sfide dell'IA senza compromettere il progresso tecnologico.

##### 5. *Conclusioni*

La regolamentazione dell'intelligenza artificiale si colloca oggi al crocevia tra diritto, tecnologia e valori fondamentali, ponendo il giurista di fronte a una sfida complessa e in continua evoluzione. Le riflessioni sviluppate nel presente contributo hanno evidenziato come gli attuali modelli normativi, seppur significativi nel tentativo di affermare *standard* etici condivisi, risultino ancora caratterizzati da una tensione irrisolta tra esigenze di tutela e promozione dell'innovazione.

Tanto la Convenzione quadro del Consiglio d'Europa quanto il Regolamento (UE) 2024/1689 testimoniano l'intento di offrire una disciplina organica del fenomeno, fondata su un approccio precauzionale e antropocentrico. Tuttavia, tali strumenti rivelano una tendenza all'universalizzazione e alla rigidità normativa che rischia di comprometterne l'effettività applicativa, soprattutto in contesti eterogenei e ad alta variabilità tecnologica. L'adozione di un'impostazione olistica, se da un lato consente di affermare principi comuni di tutela, dall'altro mostra evidenti limiti operativi: le definizioni ampie di intelligenza artificiale, l'assenza di meccanismi adattivi e la scarsa articolazione per ambiti d'uso determinano incertezza interpretativa e applicativa. Inoltre, la concentrazione sulla mitigazione del rischio *ex ante*, se non accompagnata da adeguati strumenti di governance settoriale e di cooperazione internazionale, rischia di tradursi in un effetto paralizzante sul piano dell'innovazione, in particolare per gli attori meno strutturati.

In questa prospettiva, risulta necessario superare la logica ipertrofica e orizzontale che caratterizza le attuali proposte normative, orientandosi verso modelli funzionali, settoriali e proporzionali, capaci di garantire un bilanciamento più efficace tra i diversi interessi in gioco. Una regolazione giuridica realmente adeguata al fenomeno dell'IA dovrebbe essere non solo fondata su principi, ma anche dotata di strumenti adattivi, differenziati e capaci di cogliere la specificità dei contesti applicativi.

In definitiva, la regolazione dell'intelligenza artificiale richiede un ripensamento dell'attuale impianto normativo, che, nella sua forma più recente, continua a privilegiare modelli orizzontali e generalisti. Sebbene tali approcci rispondano all'esigenza di affermare principi etici condivisi a livello internazionale, essi si rivelano, nella pratica, poco adatti a governare efficacemente la varietà e la complessità delle applicazioni tecnologiche. Una disciplina giuridica realmente efficace dovrebbe fondarsi su strumenti normativi flessibili, graduati e capaci di adattarsi alle peculiarità dei singoli contesti di impiego, senza sacrificare la coerenza del quadro complessivo. In questo senso, la sfida non è tanto quella di produrre nuove norme, quanto di affinare gli strumenti esistenti attraverso soluzioni tecniche, procedurali e istituzionali che permettano un equilibrio più solido tra tutela e innovazione.