



Freedom, Security & Justice:
European Legal Studies

Rivista giuridica di classe A

2025, n. 2

EDITORIALE
SCIENTIFICA



DIRETRICE

Angela Di Stasi

Ordinario di Diritto Internazionale e di Diritto dell'Unione europea, Università di Salerno
Titolare della Cattedra Jean Monnet 2017-2020 (Commissione europea)
"Judicial Protection of Fundamental Rights in the European Area of Freedom, Security and Justice"

CONSIGLIO SCIENTIFICO

Giandonato Caggiano, Ordinario f.r. di Diritto dell'Unione europea, Università Roma Tre
Sergio Maria Carbone, Professore Emerito, Università di Genova
Roberta Clerici, Ordinario f.r. di Diritto Internazionale privato, Università di Milano
Nigel Lowe, Professor Emeritus, University of Cardiff
Paolo Mengozzi, Professore Emerito, Università "Alma Mater Studiorum" di Bologna - già Avvocato generale presso la Corte di giustizia dell'UE
Massimo Panebianco, Professore Emerito, Università di Salerno
Nicoletta Parisi, Ordinario f.r. di Diritto Internazionale, Università di Catania - già Componente ANAC
Guido Raimondi, già Presidente della Corte EDU - già Presidente di Sezione della Corte di Cassazione
Silvana Sciarra, Professore Emerito, Università di Firenze - Presidente Emerito della Corte Costituzionale
Giuseppe Tesaurò, Professore f.r. di Diritto dell'UE, Università di Napoli "Federico II" - Presidente Emerito della Corte Costituzionale
Antonio Tizzano, Professore Emerito, Università di Roma "La Sapienza" - Vice Presidente Emerito della Corte di giustizia dell'UE
Ennio Triggiani, Professore Emerito, Università di Bari
Ugo Villani, Professore Emerito, Università di Bari

COMITATO EDITORIALE

Maria Caterina Baruffi, Ordinario di Diritto Internazionale, Università di Bergamo
Alfonso-Luis Calvo Caravaca, Catedrático Jubilado de Derecho Internacional Privado, Universidad Carlos III de Madrid
Ida Caracciolo, Ordinario di Diritto Internazionale, Università della Campania - Giudice dell'ITLOS
Pablo Antonio Fernández-Sánchez, Catedrático de Derecho Internacional, Universidad de Sevilla
Inge Govaere, Director of the European Legal Studies Department, College of Europe, Bruges
Paola Mori, Ordinario f.r. di Diritto dell'Unione europea, Università "Magna Graecia" di Catanzaro
Lina Panella, Ordinario f.r. di Diritto Internazionale, Università di Messina
Lucia Serena Rossi, Ordinario di Diritto dell'UE, Università "Alma Mater Studiorum" di Bologna - già Giudice della Corte di giustizia dell'UE



COMITATO DEI REFEREEES

Bruno Barel, Associato f.r. di Diritto dell'Unione europea, Università di Padova
Marco Benvenuti, Ordinario di Istituzioni di Diritto pubblico, Università di Roma "La Sapienza"
Francesco Buonomena, Associato di Diritto dell'Unione europea, Università di Salerno
Raffaele Cadin, Ordinario di Diritto Internazionale, Università di Roma "La Sapienza"
Ruggiero Cafari Panico, Ordinario f.r. di Diritto dell'Unione europea, Università di Milano
Federico Casolari, Ordinario di Diritto dell'Unione europea, Università "Alma Mater Studiorum" di Bologna
Luisa Cassetti, Ordinario di Istituzioni di Diritto Pubblico, Università di Perugia
Anna Cavaliere, Associato di Filosofia del diritto, Università di Salerno
Giovanni Cellamare, Ordinario f.r. di Diritto Internazionale, Università di Bari
Giuseppe D'Angelo, Ordinario di Diritto ecclesiastico e canonico, Università di Salerno
Sara De Vido, Ordinario di Diritto Internazionale, Università Ca' Foscari Venezia
Marcello Di Filippo, Ordinario di Diritto Internazionale, Università di Pisa
Rosario Espinosa Calabuig, Catedrática de Derecho Internacional Privado, Universitat de València
Valentina Faggiani, Profesora Titular de Derecho Constitucional, Universidad de Granada
Caterina Fratea, Associato di Diritto dell'Unione europea, Università di Verona
Ana C. Gallego Hernández, Profesora Ayudante de Derecho Internacional Público y Relaciones Internacionales, Universidad de Sevilla
Pietro Gargiulo, Ordinario f.r. di Diritto Internazionale, Università di Teramo
Francesca Graziani, Associato di Diritto Internazionale, Università della Campania "Luigi Vanvitelli"
Giancarlo Guarino, Ordinario f.r. di Diritto Internazionale, Università di Napoli "Federico II"
Elsbeth Guild, Associate Senior Research Fellow, CEPS
Victor Luis Gutiérrez Castillo, Profesor de Derecho Internacional Público, Universidad de Jaén
Ivan Ingravallo, Ordinario di Diritto Internazionale, Università di Bari
Paola Ivaldi, Ordinario di Diritto Internazionale, Università di Genova
Luigi Kalb, Ordinario di Procedura Penale, Università di Salerno
Luisa Marin, Ricamatore di Diritto dell'UE, Università dell'Insubria
Simone Marinai, Associato di Diritto dell'Unione europea, Università di Pisa
Fabrizio Marongiu Buonaiuti, Ordinario di Diritto Internazionale, Università di Macerata
Rostane Medhi, Professeur de Droit Public, Université d'Aix-Marseille
Michele Messina, Ordinario di Diritto dell'Unione europea, Università di Messina
Stefano Montaldo, Associato di Diritto dell'Unione europea, Università di Torino
Violeta Moreno-Lax, Senior Lecturer in Law, Queen Mary University of London
Claudia Morviducci, Professore Senior di Diritto dell'Unione europea, Università Roma Tre
Michele Nino, Ordinario di Diritto Internazionale, Università di Salerno
Criseide Novi, Associato di Diritto Internazionale, Università di Foggia
Anna Oriolo, Associato di Diritto Internazionale, Università di Salerno
Leonardo Pasquali, Ordinario di Diritto internazionale, Università di Pisa
Piero Pennetta, Ordinario f.r. di Diritto Internazionale, Università di Salerno
Francesca Perrini, Associato di Diritto Internazionale, Università di Messina
Gisella Pignataro, Associato di Diritto privato comparato, Università di Salerno
Emanuela Pistoia, Ordinario di Diritto dell'Unione europea, Università di Teramo
Anna Pitrone, Associato di Diritto dell'Unione europea, Università di Messina
Concetta Maria Pontecorvo, Ordinario di Diritto Internazionale, Università di Napoli "Federico II"
Pietro Pustorino, Ordinario di Diritto Internazionale, Università LUISS di Roma
Santiago Ripol Carulla, Catedrático de Derecho internacional público, Universitat Pompeu Fabra Barcelona
Angela Maria Romito, Associato di Diritto dell'Unione europea, Università di Bari
Gianpaolo Maria Ruotolo, Ordinario di Diritto Internazionale, Università di Foggia
Teresa Russo, Associato di Diritto dell'Unione europea, Università di Salerno
Alessandra A. Souza Silveira, Diretora do Centro de Estudos em Direito da UE, Universidad do Minho
Ángel Tinoco Pastrana, Profesor de Derecho Procesal, Universidad de Sevilla
Sara Tonolo, Ordinario di Diritto Internazionale, Università degli Studi di Padova
Chiara Enrica Tuo, Ordinario di Diritto dell'Unione europea, Università di Genova
Talitha Vassalli di Dachenhausen, Ordinario f.r. di Diritto Internazionale, Università di Napoli "Federico II"
Valentina Zambrano, Associato di Diritto Internazionale, Università di Roma "La Sapienza"
Alessandra Zanobetti, Ordinario f.r. di Diritto Internazionale, Università "Alma Mater Studiorum" di Bologna

COMITATO DI REDAZIONE

Angela Festa, Docente incaricato di Diritto dell'Unione europea, Università della Campania "Luigi Vanvitelli"
Anna Iermano, Associato di Diritto Internazionale, Università di Salerno
Daniela Marrani, Associato di Diritto Internazionale, Università di Salerno
Rossana Palladino (Coordinatore), Associato di Diritto dell'Unione europea, Università di Salerno

Revisione linguistica degli abstracts a cura di

Francesco Campofreda, Dottore di ricerca in Diritto Internazionale, Università di Salerno



Rivista quadrimestrale on line "Freedom, Security & Justice: European Legal Studies" www.fsjeurostudies.eu
Editoriale Scientifica, Via San Biagio dei Librai, 39 - Napoli

CODICE ISSN 2532-2079 - Registrazione presso il Tribunale di Nocera Inferiore n° 3 del 3 marzo 2017



Indice-Sommario 2025, n. 2

Editoriale

Dalla dichiarazione Schuman al Libro bianco sulla prontezza alla difesa europea: verso una revisione del progetto europeo? p. 1
Ugo Villani

Saggi, Articoli, Commenti e Note

Le origini dello Spazio di libertà, sicurezza e giustizia. Pace e conflitti armati (1945-2025) p. 14
Massimo Panebianco

Migrare: un diritto fondamentale? p. 26
Antonio Ruggeri

Il ruolo della Procura europea (EPPO) nella tutela dello Stato di diritto dell'Unione europea p. 42
Serena Crespi

Norme di diritto internazionale e disparità di genere, idee vecchie e nuove. Il caso del *mundio muliebre*, uno stereotipo da rileggere p. 82
Lucia di Cintio

Convenzione delle Nazioni Unite contro il *cybercrime* e tutela dei diritti umani: influenze europee sullo scenario internazionale p. 108
Marco Dimetto

The error in predictive justice systems. Challenges for justice, freedom, and human-centrism p. 131
under EU law
Alessandro Ferrara

EU impact on Albanian medical civil liability: a case law approach p. 146
Enkelejda Koka, Denard Veshi, Aisha Morina

La promozione della parità di genere nelle relazioni tra l'Unione europea e i *partner* meridionali p. 162
Claudia Morini



FOCUS

Democracy and the Rule of Law: A New Push for European Values

Il Focus contiene contributi elaborati a seguito della riflessione realizzata nel Seminario conclusivo dello Jean Monnet Module Eu-Draw (2022-2025) "Democracy and the Rule of Law: A New Push for European Values", tenutosi presso l'Università degli Studi di Salerno (1 aprile 2025)

- Presentazione del *Focus* p. 192
Angela Di Stasi
- Values in the EU external action: mechanisms of implementation and their outcomes p. 194
Stefania Kolarz
- Justice and Home Affairs Cooperation (JHAC) in the perspective of enlargement p. 211
Teresa Russo
- Brevi riflessioni sulla tutela dei diritti nello "spazio digitale" europeo p. 228
Francesco Buonomenna
- Consiglio d'Europa e intelligenza artificiale: un primo tentativo di regolamentazione a tutela dei diritti umani, democrazia e Stato di diritto p. 242
Anna Iermano
- La disinformazione *online* come "minaccia ibrida" alla democrazia nell'Unione europea: meccanismi di tutela e strumenti a contrasto per uno Spazio di libertà, sicurezza e giustizia p. 272
Angela Festa
- L'"approccio europeo" al contrasto alla disinformazione digitale e alla protezione dei valori democratici: quale contributo dell'*AI Act*? p. 296
Rossana Palladino



CONSIGLIO D'EUROPA E INTELLIGENZA ARTIFICIALE: UN PRIMO TENTATIVO DI REGOLAMENTAZIONE A TUTELA DI DIRITTI UMANI, DEMOCRAZIA E STATO DI DIRITTO

Anna Iermano *

SOMMARIO: 1. Premessa. – 2. Introduzione alla genesi della Convenzione “quadro” sull’intelligenza artificiale e i diritti umani, la democrazia e lo Stato di diritto. – 3. Obiettivi della Convenzione quadro e modalità attuative. – 4. Ambito applicativo delle disposizioni convenzionali e limiti. – 5. Obblighi generali e principi a garanzia di sistemi di intelligenza artificiale nel rispetto di diritti umani, democrazia e *rule of law*. – 6. Rimedi e gestione dei rischi e degli impatti negativi. – 7. A proposito del meccanismo di *follow-up* e della cooperazione. – 8. La partecipazione attiva dell’Unione europea ai negoziati della Convenzione quadro: spunti di riflessione. – 9. Considerazioni conclusive sul *quid pluris* della Convenzione quadro.

1. Premessa

Come noto, gli Stati-parte del Consiglio d’Europa si sono impegnati a garantire i diritti e le libertà sanciti dalla Convenzione europea per la salvaguardia dei diritti dell’uomo e delle libertà fondamentali (CEDU) a chiunque rientri nella loro giurisdizione e, per quanto qui rileva, tale impegno permane anche a fronte del progresso tecnologico e della trasformazione digitale in corso¹.

Ne discende la responsabilità di assicurare che qualsiasi progettazione, sviluppo e implementazione continua di sistemi algoritmici avvenga nel rispetto dei diritti umani e delle libertà fondamentali, che sono universali, indivisibili, interdipendenti e interconnessi, al fine di amplificare gli effetti positivi e prevenire o minimizzare i possibili risvolti negativi.

Le attività rientranti nel ciclo di vita dei sistemi di intelligenza artificiale presentano, infatti, al contempo, vantaggi e difetti, specie in punto di tutela dei diritti umani,

Double-blind peer reviewed article.

* Professore associato di Diritto Internazionale, Università degli Studi di Salerno. Indirizzo e-mail: aiermano@unisa.it.

¹ Recommendation of the Committee of Ministers to member States on the human rights impacts of algorithmic systems (Adopted by the Committee of Ministers on 8 April 2020 at the 1373rd meeting of the Ministers’ Deputies), CM/Rec(2020)1.

democrazia e Stato di diritto²: per un verso esse hanno il pregio di promuovere la prosperità umana e il benessere delle persone e della società, lo sviluppo sostenibile, l'uguaglianza di genere e l'emancipazione delle donne³, rafforzando il progresso e l'innovazione, e offrendo opportunità senza precedenti per salvaguardare e promuovere i diritti umani, la democrazia e la *rule of law*. In particolare, esse posseggono un potenziale significativo per migliorare le istituzioni e i processi democratici, consentendo una maggiore partecipazione dei cittadini al dibattito pubblico, così come una condivisione diffusa delle informazioni nello spazio civico *online*, garantendo maggiore inclusività, trasparenza ed efficienza.

Per altro verso, talune attività possono pregiudicare la dignità umana e l'autonomia individuale, il rispetto dei diritti umani, l'osservanza dello Stato di diritto, come pure possono rappresentare una minaccia per la democrazia, influenzare in modo occulto o manipolativo le scelte individuali, le opinioni o le decisioni e minare così la partecipazione civica e la fiducia collettiva.

Esse, ad esempio, possono consentire agli utenti, specie se malintenzionati, di diffondere disinformazione e/o cattiva informazione, tramite l'uso di contenuti generati dall'IA o la manipolazione, da parte della stessa, di contenuti autentici; generare decisioni pregiudizievoli sugli individui, con conseguenti pratiche discriminatorie; portare all'esclusione di taluni o di coloro che sono sottorappresentati o che versano in una situazione di vulnerabilità; comportare restrizioni della libertà di riunione o di espressione e/o della privacy; influenzare le sentenze dei tribunali, con potenziali implicazioni sull'integrità del sistema giudiziario, sì da indurre la Commissione europea per l'efficienza della giustizia (CEPEJ), già nel 2018, all'adozione della Carta etica europea sull'utilizzo dell'intelligenza artificiale nei sistemi giudiziari e negli ambiti connessi⁴.

Si pensi, inoltre, *inter alia*, ai rischi di discriminazione nei contesti digitali, in quanto attività potenzialmente idonee a creare o aggravare le disuguaglianze, comprese quelle subite da donne, in relazione all'esercizio dei loro diritti e alla partecipazione piena,

² Preambolo della Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, Vilnius, 5.IX.2024.

³ Sul punto si segnala A. DI STASI (a cura di), *Dalla non discriminazione alle pari opportunità. Intelligenza Artificiale e nuove Tecnologie, Sostenibilità e Gender Gap: il ruolo dell'Università*. Atti del Convegno di studi tenuto in occasione della Giornata internazionale della donna 2024 - Università degli Studi di Salerno, Milano, 2025.

⁴ Carta etica europea sull'utilizzo dell'intelligenza artificiale nei sistemi giudiziari e negli ambiti connessi adottata dalla Commissione europea per l'efficienza della giustizia (CEPEJ), Strasburgo 3-4 dicembre 2018, nel corso della sua 31^a Riunione plenaria, <https://rm.coe.int/carta-etica-europea-sull-utilizzo-dell-intelligenza-artificiale-nei-si/1680993348>. Tale Carta fissa cinque principi: 1) principio del rispetto dei diritti fondamentali: assicurare l'elaborazione e l'attuazione di strumenti e servizi di intelligenza artificiale siano compatibili con i diritti fondamentali; 2) principio di non-discriminazione: prevenire specificamente lo sviluppo o l'intensificazione di discriminazioni tra persone o gruppi di persone; 3) principio di qualità e sicurezza: in ordine al trattamento di decisioni e dati giudiziari, utilizzare fonti certificate e dati intangibili con modelli elaborati multidisciplinariamente, in un ambiente tecnologico sicuro; 4) principio di trasparenza, imparzialità ed equità: rendere le metodologie di trattamento dei dati accessibili e comprensibili, autorizzare verifiche esterne; 5) principio "del controllo da parte dell'utilizzatore": precludere un approccio prescrittivo e assicurare che gli utilizzatori siano attori informati e abbiano il controllo delle loro scelte.

paritaria ed efficace agli affari economici, sociali, culturali e politici⁵; o ai rischi di un uso improprio dei sistemi di intelligenza artificiale a fini repressivi, in violazione del diritto internazionale dei diritti umani, anche attraverso pratiche di sorveglianza e censura arbitrarie e illegittime che minano la tutela della vita privata e l'autonomia.

Da qui la necessità di contemperare pregi e difetti, prevenendo e limitando la possibile compromissione di diritti umani, democrazia e Stato di diritto, senza per questo rinunciare alle crescenti potenzialità insite nei sistemi di intelligenza artificiale⁶.

Ed è sulla base di tali premesse che nasce la “Convenzione quadro⁷ sull'intelligenza artificiale e i diritti umani, la democrazia e lo Stato di diritto”, adottata il 17 maggio 2024 dal Comitato dei Ministri del Consiglio d'Europa⁸ nella sua 133^a sessione tenutasi a Strasburgo e aperta alla firma in occasione della Conferenza dei Ministri della giustizia degli Stati-parte del Consiglio d'Europa a Vilnius (Lituania) il 5 settembre 2024.

Così premesso, in questa sede si intende fornire un'analisi critica e puntuale dei contenuti di siffatta Convenzione, evidenziando talora profili di affinità o divergenza rispetto al coevo Regolamento (UE) 2024/1689 che stabilisce regole armonizzate sull'intelligenza artificiale⁹, meglio conosciuto come *AI Act (Artificial Intelligence Act)*, per poi concludere con qualche riflessione sul *quid pluris* del testo convenzionale.

⁵ C. CANTILLO, *Intelligenza Artificiale e questioni di genere: alcune considerazioni*, in A. DI STASI (a cura di), *Dalla non discriminazione alle pari opportunità*, cit., p. 186. Come osserva l'A., se si tratta senza dubbio di una “trasformazione” – per molti aspetti di una “rivoluzione”, avvenuta in seguito all'avvento del digitale – non si può con altrettanta certezza sostenere che questa trasformazione sia unidirezionalmente orientata “verso il meglio”. E come auspica l'A. (p. 187), occorre salvaguardare la dimensione della persona, scongiurando il pericolo che sempre più sia la nostra intelligenza a “macchinizzarsi” fino a ibridarsi con la macchina, o, ancora oltre, che le macchine intelligenti nelle loro capacità generative possano sopravanzare e prendere il posto degli umani, con esiti incontrollabili.

⁶ Nell'ambito dell'Unione europea, nel Libro bianco sull'intelligenza artificiale – Un approccio europeo all'eccellenza e alla fiducia, COM (2020) 65final, del 19 febbraio 2020, la Commissione aveva evidenziato l'intento di mantenere il delicato equilibrio tra l'esigenza di «preservare la leadership tecnologica dell'UE», incoraggiando le imprese a sviluppare le soluzioni basate sull'IA», e la necessità di assicurare «che i cittadini europei possano beneficiare di nuove tecnologie sviluppate e operanti in conformità ai valori, ai diritti fondamentali e ai principi dell'Unione».

⁷ Il carattere di Convenzione “quadro” è esplicitamente asserito nel Preambolo, al considerando 11: «Recognising the framework character of this Convention, which may be supplemented by further instruments to address specific issues relating to the activities within the lifecycle of artificial intelligence systems». Il “*framework character*” non esclude, dunque, la possibile adozione di ulteriori strumenti per affrontare questioni specifiche relative alle attività che rientrano nel ciclo di vita dei sistemi di intelligenza artificiale. Essa può essere integrata tramite protocolli aggiuntivi che dovrebbero incoraggiare la valutazione dei rischi associati a tali tecnologie, inclusi quelli relativi alla salute umana, all'ambiente e agli aspetti socioeconomici, come lavoro e occupazione.

⁸ Explanatory Report to the Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, Vilnius, 5.IX.2024.

⁹ 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), PE/24/2024/REV/1, in GU L 2024/1689, 12.7.2024, il quale contiene norme armonizzate basate in linea generale sulla piena armonizzazione che disciplinano l'immissione sul mercato, la messa in servizio e l'uso di sistemi di IA nell'Unione. Cfr., N. ZORZI GALGANO, *Il Regolamento dell'Unione Europea del 2024 sull'intelligenza artificiale. Una prospettiva di sistema*, in *Contratto e impresa. Europa (Online)*, 2025, fasc. 1, p. 1 ss.

2. Introduzione alla genesi della Convenzione “quadro” del Consiglio d’Europa sull’intelligenza artificiale e i diritti umani, la democrazia e lo Stato di diritto

Con la firma da parte dei primi paesi (Andorra, Georgia, Islanda, Norvegia, Repubblica di Moldavia, San Marino, Regno Unito, Israele, Stati Uniti d’America¹⁰)¹¹ e dell’Unione europea, la Convenzione rappresenta il primo trattato atto a vincolare gli Stati a prevenire e mitigare i rischi associati all’uso dell’intelligenza artificiale, in un contesto internazionale.

Aperta alla sottoscrizione pure di Paesi che non hanno partecipato alle negoziazioni (art. 31, par. 1 della Convenzione)¹², anche non europei¹³, essa entrerà in vigore dopo la ratifica da parte di cinque sottoscrittori¹⁴, inclusi almeno tre Stati-parte del COE, ovvero all’indomani di un numero esiguo di ratifiche, il che è dovuto probabilmente – come si specificherà a breve – ad una negoziazione piuttosto partecipativa e inclusiva¹⁵, tale da

¹⁰ Si segnala USA, *International Cyberspace and Digital Policy Strategy: Towards an Innovative, Secure, and Rights-Respecting Digital Future*, pubblicata dal Dipartimento di Stato americano il 6 maggio 2024, <https://www.state.gov/united-states-international-cyberspace-and-digitalpolicy-strategy/>. Si tratta di una strategia elaborata dal Dipartimento di Stato, con il supporto di altre agenzie federali, per fungere da guida nella sua azione internazionale in materia di diplomazia tecnologica e far progredire la *National Security Strategy* come pure la *National Cybersecurity Strategy*. La strategia è incentrata sul concetto di solidarietà digitale, ovvero la volontà di lavorare insieme su obiettivi condivisi, stare uniti, aiutare i partner a sviluppare capacità e fornire supporto reciproco. Cfr. F. ZORZI GIUSTINIANI, *La Convenzione quadro sull’intelligenza artificiale e i diritti umani, la democrazia e lo stato di diritto e l’International Cyberspace and Digital Policy Strategy statunitense*, in *Nomos*, 2024, n. 2, p. 1 ss., <https://www.nomos-leattualitaneldiritto.it/wp-content/uploads/2024/10/zorzigiustiniani.cronache2-2024.pdf>.

¹¹ Ad essi si sono aggiunti il Canada e il Giappone che hanno firmato la Convenzione quadro. La cerimonia di firma si è tenuta prima dell’evento del Consiglio d’Europa incentrato sul coinvolgimento dell’Africa nella governance mondiale dell’IA e organizzato a margine del Vertice per l’azione sull’IA, ospitato dalla Francia. Hanno firmato, altresì, Liechtenstein, Montenegro, Svizzera, Ucraina. I firmatari sono consultabili al seguente indirizzo: <https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence>.

¹² Vedi, ad esempio, Convenzione del Consiglio d’Europa sul riciclaggio, la ricerca, il sequestro e la confisca dei proventi di reato e sul finanziamento del terrorismo (STCE no. 198), anch’essa aperta alla firma pure di Stati non parte del Consiglio d’Europa. In tal caso fu firmata e ratificata solo dal Marocco.

¹³ C. NARDOCCI, *La (seconda) svolta del 2024. Anche il Consiglio d’Europa decide di regolamentare l’intelligenza artificiale*, in *BioLaw Journal – Special issue*, 2024, n. 1, p. 73 ss. L’A. circa l’estensione della Convenzione oltre i confini del continente europeo, suggerisce almeno tre considerazioni. La prima è che persiste una tendenza degli Stati a preferire che siano organizzazioni internazionali a disciplinare un fenomeno di difficile contenimento su scala nazionale; la seconda è che simile preferenza per la dimensione sovrastatale ha spinto ordinamenti giuridici, per lungo tempo scettici nei confronti di una regolamentazione delle tecniche di intelligenza artificiale, a voltare il proprio sguardo verso chi quella direzione ha deciso, invece, di intraprendere (es. Stati Uniti, Canada); in terzo luogo, la centralità della *Framework Convention* va inquadrata anche in prospettiva orizzontale e non solo verticale e, cioè, nelle sue relazioni con il dato statale.

¹⁴ La Convenzione entrerà in vigore il primo giorno del mese successivo alla scadenza di un periodo di tre mesi dalla data in cui i cinque firmatari, tra cui almeno tre Stati membri del Consiglio d’Europa, l’avranno ratificata.

¹⁵ In senso critico, invece, M. CASTELLANETA, *Al via il primo trattato globale sull’intelligenza artificiale*, 7 giugno 2024, <https://www.affarinternazionali.it/al-via-il-primo-trattato-globale-sullintelligenza-artificiale-2/>. L’A. sostiene che, per l’entrata in vigore della Convenzione sia richiesto un numero basso di ratifiche, elemento, questo, che rischia di intaccare l’efficacia del Trattato e che mostra più un desiderio di avere un trattato quale che sia, rispetto alla volontà di incidere in modo effettivo sulla regolamentazione.

tradursi “a monte” in un accordo ampiamente condiviso, di cui si auspica la vincolatività il prima possibile.

Il testo, che si compone di un preambolo e di trentasei articoli, è stato redatto, infatti, dopo oltre 4 anni di lavori¹⁶, dal Comitato sull'intelligenza artificiale (CAI)¹⁷, quale organismo intergovernativo istituito dal Consiglio d'Europa che ha riunito i 46 Stati parte del COE e 11 Stati non membri¹⁸ che condividono i valori e gli obiettivi del Consiglio d'Europa stesso (Argentina, Australia, Canada, Costa Rica, Giappone, Israele, Messico, Perù, Santa Sede, Stati Uniti d'America e Uruguay), oltre 68 *stakeholders*, rappresentanti della società civile, del settore privato, e del mondo accademico, che hanno partecipato in qualità di osservatori. Né la Russia, né la Bielorussia sono state rappresentate nel Comitato a causa, come noto, dell'invasione dell'Ucraina da parte dell'una e della partecipazione attiva all'invasione da parte dell'altra.

Tale Comitato ha così basato il suo lavoro di stesura del testo¹⁹, come anticipato, su un processo di consultazione notevolmente inclusivo e trasversale, anche con partner internazionali e sovranazionali, quali: l'Organizzazione per la sicurezza e la cooperazione

¹⁶ I punti chiave della riflessione del Consiglio d'Europa su intelligenza artificiale e diritti fondamentali a partire da ottobre 2021 sono disponibili al seguente indirizzo: <https://www.coe.int/en/web/artificial-intelligence/work-in-progress#02EN>.

¹⁷ Il *Committee on Artificial Intelligence (CAI)*, istituito nel gennaio 2022 è succeduto all'*Ad Hoc Committee on Artificial Intelligence (CAHAI)*, istituito l'11 settembre 2019. Attraverso il lavoro dei comitati il Consiglio d'Europa si occupa da tempo dei problemi che l'umanità deve affrontare a causa dei progressi nelle tecnologie dell'informazione e digitali e, in particolare, nei sistemi algoritmici e di intelligenza artificiale. Si veda il *final paper* adottato dal CAHAI “*Possible elements of a legal framework on artificial intelligence, based on Council of Europe’s standards on human rights, democracy and the rule of law 2021*”, dicembre 2021. Sulla base di tale *final report* il Comitato dei ministri del Consiglio d'Europa ha incaricato il CAI di elaborare una Convenzione quadro sulle attività nell'ambito del ciclo di vita dei sistemi di intelligenza artificiale «based on the Council of Europe’s standards on human rights, democracy and the rule of law, and conducive to innovation, which can be composed of a binding legal instrument of a transversal character, including notably general common principles» (così Report, cit., *Introduction*, p. 1).

¹⁸ Hanno aderito al processo di negoziazione nel CAI e hanno partecipato all'elaborazione della presente Convenzione quadro in qualità di Stati osservatori.

¹⁹ Nel corso della negoziazione e della successiva adozione della Convenzione quadro, sono stati presi in considerazione i seguenti strumenti giuridici e politici internazionali sull'intelligenza artificiale, in particolare quelli predisposti dal Consiglio d'Europa, ma anche da altre organizzazioni internazionali: Declaration of the Committee of Ministers of the Council of Europe on the manipulative capabilities of algorithmic processes, adopted on 13 February 2019 - Decl(13/02/2019); Recommendation on Artificial Intelligence adopted by the OECD Council on 22 May 2019 (the “OECD AI Principles”); Recommendation of the Committee of Ministers of the Council of Europe to member States on the human rights impacts of algorithmic systems, adopted on 8 April 2020 - CM/Rec(2020)1; Resolutions and Recommendations of the Parliamentary Assembly of the Council of Europe, examining the opportunities and risks of artificial intelligence for human rights, democracy, and the rule of law and endorsing a set of core ethical principles that should be applied to AI systems; UNESCO Recommendation on the Ethics of Artificial Intelligence adopted on 23 November 2021; G7 Hiroshima Process International Guiding Principles for Organisations Developing Advanced AI Systems and Hiroshima Process International Code of Conduct for Organizations Developing Advanced AI Systems (adopted on 30 October 2023); and EU Regulation laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) adopted on 13 March 2024. Inoltre, i negoziati si sono basati anche sulle seguenti dichiarazioni politiche: Declaration by Heads of State and Government made at the 4th Council of Europe Summit in Reykjavik on 16-17 May 2023; G7 Leaders’ Statement on the Hiroshima AI Process of 30 October and 6 December 2023; The Bletchley Declaration by Countries Attending the AI Safety Summit, 1-2 November 2023.

in Europa (OSCE), l'Organizzazione per la cooperazione e lo sviluppo economico (OCSE), l'Organizzazione delle Nazioni Unite per l'educazione, la scienza e la cultura (UNESCO), e ciò al fine di garantire una visione globale e condivisa sul tema.

Anche l'Unione europea – come si approfondirà in seguito – ha preso parte alle negoziazioni tramite la Commissione UE, in virtù della decisione 2022/2349²⁰ e includendo nella sua delegazione rappresentanti dell'Agenzia dell'Unione europea per i diritti fondamentali (FRA)²¹ e del Garante europeo della protezione dei dati (EDPS).

Degno di nota è, pertanto e, anzitutto, l'approccio multilivello (nazionale, europeo, internazionale), multidisciplinare e plurisoggettivo che ha contraddistinto i negoziati, coinvolgendo, tra gli altri, civili, privati, accademici, per far fronte in maniera efficace ed efficiente a sfide e opportunità poste dall'IA; un *iter* “orizzontale” che, lungi dall'imporre “dall'alto” una regolamentazione “quadro” nel caso dell'IA, interpella gli stessi soggetti interessati, offrendo, altresì, l'occasione per una riflessione sulla scala valoriale comune nell'ottica della condivisione e del mutuo confronto.

3. Obiettivi della Convenzione quadro e modalità attuative

La Convenzione *de qua*, così negoziata, una volta in vigore, costituirà il primo accordo internazionale sull'IA che vincola le parti aderenti a garantire che le attività rientranti nell'intero *life cycle* dei sistemi di IA, dalla progettazione alla dismissione, da parte di attori sia pubblici che privati, siano pienamente conformi ai diritti umani – come del resto evidenziato anche dall'Assemblea Generale delle Nazioni Unite²², sia pure in un atto di *soft law*²³ –, alla democrazia e allo Stato di diritto (art. 1), riducendo al minimo il rischio che questi siano compromessi.

²⁰ Con la decisione (UE) 2022/2349 del Consiglio, del 21 novembre 2022, che autorizza l'avvio di negoziati a nome dell'Unione europea per una convenzione del Consiglio d'Europa sull'intelligenza artificiale, i diritti umani, la democrazia e lo Stato di diritto, in GU L 311 del 2 dicembre 2022, p. 138, il Consiglio ha autorizzato la Commissione ad avviare negoziati a nome dell'Unione.

²¹ In ambito UE si segnala la relazione dell'Agenzia dell'Unione europea per i diritti fondamentali (FRA), *Preparare un giusto futuro - L'intelligenza artificiale e i diritti fondamentali nell'UE*, Lussemburgo, 2021, che fornisce un'istantanea dell'uso attuale delle tecnologie legate all'IA nell'Unione europea (UE) e ne analizza le implicazioni per i diritti fondamentali. La relazione si concentra su casi d'uso in quattro ambiti fondamentali: prestazioni sociali, polizia predittiva, servizi sanitari e pubblicità mirata.

²² Come si precisa nel *report* l'espressione “diritti umani” utilizzata nella presente Convenzione quadro intende includere le “libertà fondamentali” nelle giurisdizioni che tracciano una distinzione tra i due termini.

²³ Risoluzione dell'Assemblea Generale delle Nazioni Unite del 11 marzo 2024, *Seizing the opportunities of safe, secure and trustworthy artificial intelligence systems for sustainable development*, <https://docs.un.org/en/A/78/L.49> in cui l'Assemblea sottolinea il rispetto, la protezione e la promozione dei diritti umani nella progettazione, nello sviluppo, nell'implementazione e nell'utilizzo dell'IA. Inoltre, menziona anch'essa, sin dal preambolo, quale ambito di applicazione l'intero “ciclo di vita” (*life cycle*) dell'intelligenza artificiale e lega, per la prima volta, il tema dello sviluppo sostenibile all'intelligenza artificiale (§ 3, punto d). Nel contesto dell'Organizzazione delle Nazioni Unite, si segnala, l'istituzione, nel 2023, di un *High-Level Advisory Body on Artificial Intelligence*: <https://www.un.org/digital-emerging-technologies/ai-advisory-body>.

Essa è incentrata sulla tutela e promozione dei diritti umani, della democrazia²⁴ e della *rule of law* e non regola espressamente gli aspetti economici e di mercato dei sistemi di intelligenza artificiale²⁵, né le tecnologie di IA in quanto tali.

Oggetto e finalità della Convenzione sono, pertanto, limitate a questioni attinenti al mandato del Consiglio d'Europa: essa, nel suo complesso, fornisce un quadro giuridico comune a livello globale, al fine di garantire che gli obblighi giuridici internazionali e nazionali esistenti, che incombono su ciascuna Parte nella sfera dei diritti umani, della democrazia e dello Stato di diritto, si applichino anche rispetto all'IA.

Ebbene, l'estensione di tali obblighi, già esistenti, all'intero *lifecycle* dei sistemi di IA – senza, dunque, specificare le attività rilevanti²⁶ – garantisce che la Convenzione possa coprire tutte le fasi, facendo fronte non solo ai rischi attuali ma anche quelli futuri, attesa la rapidità e spesso imprevedibilità degli sviluppi tecnologici.

Come si legge nel Preambolo, l'obiettivo fondamentale della Convenzione quadro è garantire che il potenziale delle tecnologie di intelligenza artificiale per promuovere la prosperità umana, il benessere individuale e sociale e per rendere il mondo più produttivo, innovativo e sicuro, sia sfruttato in maniera responsabile, rispetti, protegga e realizzi i valori condivisi delle Parti e sia ossequioso dei diritti umani, della democrazia e dello Stato di diritto, in ciò discostandosi, in parte, dall'*AI Act* ove, nel Preambolo e all'art. 1, si dichiara, anzitutto, che scopo del regolamento è migliorare il funzionamento del mercato interno, nella specie di mercato unico digitale²⁷, come d'altronde si evince dalla

²⁴ G. ZACCARONI, *Intelligenza artificiale e principio democratico: riflessioni a margine dell'emersione di un quadro normativo europeo*, in *Quaderni AISDUE*, Fascicolo speciale 2024, n. 2, p. 12, ove l'A. evidenzia che tale Convenzione è il primo strumento di diritto internazionale ad occuparsi dell'impatto delle nuove tecnologie sull'ordinamento democratico.

²⁵ M. CASTELLANETA, *Al via il primo trattato globale sull'intelligenza artificiale*, cit. A parere dell'A., pur trattandosi di un trattato globale unico, appare chiaro che, anche per la genericità degli obblighi e per lo spazio lasciato a interventi statali, non sarà in grado di incidere sulle multinazionali della tecnologia come il regolamento Ue sull'intelligenza artificiale (*IA Act*), tanto più che la Convenzione non riguarda gli aspetti economici e di mercato.

²⁶ La Convenzione non fornisce un elenco esaustivo di attività all'interno del ciclo di vita che sono specifiche per i sistemi di intelligenza artificiale, perché mirano a coprire tutte le attività, dalla progettazione di un sistema di intelligenza artificiale al suo ritiro, indipendentemente dall'attore coinvolto. È scelta intenzionale dei redattori di non specificarle esplicitamente poiché potrebbero dipendere dal tipo di tecnologia e da altri elementi contestuali e cambiare nel tempo, traendo ispirazione dall'ultimo lavoro dell'OCSE, al momento della stesura, i redattori forniscono come esempi di attività rilevanti: (1) pianificazione e progettazione, (2) raccolta ed elaborazione dei dati, (3) sviluppo di sistemi di intelligenza artificiale, inclusa la creazione di modelli e/o la messa a punto di modelli esistenti per compiti specifici, (4) test, verifica e convalida, (5) fornitura/messa a disposizione dei sistemi per l'uso, (6) distribuzione, (7) funzionamento e monitoraggio e (8) ritiro.

²⁷ Comunicazione della Commissione al Parlamento europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle Regioni, *Strategia per il mercato unico digitale in Europa* (MUD), COM(2015)192 final, del 6 maggio 2015. Circa l'ingresso dell'IA nel contesto del MUD, cfr. 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 dell'intelligenza artificiale*, Pisa, 2020. Vedi, altresì, Comunicazione della Commissione al Parlamento europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle Regioni, *L'intelligenza artificiale per l'Europa*, COM (2018) 237final, del 25 aprile 2018; Libro bianco sull'intelligenza artificiale, cit.; Comunicazione della Commissione al Parlamento europeo,

sua base giuridica, l'art. 114 del Trattato sul funzionamento dell'Unione europea (TFUE), cui si affianca l'art. 16 TFUE in tema di protezione e circolazione dei dati.

Allo stesso tempo essa intende promuovere lo sviluppo dell'IA attraverso un'innovazione sicura – così consacrando il cd. principio della *safe innovation* – nonché antropocentrica e responsabile, sotto la supervisione di autorità competenti (art. 13). Quanto al principio appena menzionato, trattasi di criterio guida innovativo che deriva dall'applicazione del principio di precauzione all'ambito dell'innovazione tecnologica (vedi le *Regulatory Sandboxes*)²⁸, e che consente di evitare, ad esempio, l'automazione di decisioni pubbliche e private, che possa condurre a discriminazioni sistemiche o all'erosione della *privacy*.

Del pari anche il regolamento europeo, ancorché miri ad uno scopo diverso, intende promuovere la diffusione di un'intelligenza artificiale antropocentrica e affidabile, garantendo nel contempo un livello elevato di protezione della salute, della sicurezza e dei diritti fondamentali sanciti dalla Carta di Nizza/Strasburgo, compresi la democrazia, lo Stato di diritto e la protezione dell'ambiente²⁹ – non citata, invece, nella Convenzione quadro³⁰ – contro gli effetti nocivi dei sistemi di IA nell'Unione, e promuovendo l'innovazione (art. 1 e considerando n. 1).

Tornando alla Convenzione viene, quindi, richiesto a ogni Parte firmataria di “adottare o mantenere in vigore” misure legislative, amministrative o di altra natura, anche non giuridicamente vincolanti (ad es. orientamenti interpretativi, circolari,...) che siano adeguate a dare attuazione alla Convenzione medesima (art. 1, par. 2), nonché

al Consiglio, al Comitato economico e sociale europeo e al Comitato delle Regioni, *Promuovere un approccio europeo all'intelligenza artificiale*, COM(2021)205 final, del 21 aprile 2021.

²⁸ Vedi, ad esempio le *Regulatory Sandboxes* che sono ambienti controllati dove aziende e istituzioni possono testare nuove tecnologie come l'intelligenza artificiale, prima di una regolamentazione definitiva, in modo da valutare rischi e benefici in un quadro giuridico delimitato. In questo contesto, l'intelligenza artificiale può essere sperimentata in modo sicuro, con la supervisione delle autorità competenti che monitorano il funzionamento delle tecnologie e intervengono per correggere eventuali criticità. Al riguardo, G. RUGANI, *La promozione di strumenti di co-regolazione dell'intelligenza artificiale nell'AI Act, con particolare riferimento alle regulatory sandboxes*, in *Quaderni AISDUE*, fasc. speciale 2024, n. 2, p. 143 ss., <https://www.aisdue.eu/wp-content/uploads/2024/10/AI-2-24-spec.pdf>.

²⁹ Vedi emendamenti del Parlamento europeo, approvati il 14 giugno 2023, alla 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 (P9_TA(2023)0236).

³⁰ Giova rammentare che neppure nella CEDU esiste una norma che consacra la tutela ambientale, atteso che nel 1950 la sensibilità ambientale era praticamente inesistente. Il testo convenzionale contempla solo i cd. diritti di prima generazione (quelli civili e politici) e, dunque, non poteva contemplare un diritto cd. di nuova generazione espressivo di una coscienza collettiva all'epoca inesistente. Tuttavia, la Corte europea, nonostante l'assenza di una norma, ha utilizzato gli articoli 2 (tutela del diritto alla vita), 3 (tutela integrità psico-fisica), 8 (tutela della vita privata e familiare) e 1 del Protocollo n. 1 (tutela della proprietà) per estendere la tutela anche all'ambiente (es. Corte europea dei diritti dell'uomo, sentenza del 10 gennaio 2012, ricorso n. 30765/08, *Di Sarno ed altri c. Italia*; sentenza del 24 gennaio 2019, ricorsi nn. 54414/13 e 54264/15, *Cordella e altri c. Italia*; Grande Camera, sentenza del 9 aprile 2024, ricorso n. 53600/20, *Verein KlimaSeniorinnen Schweiz e altri c. Svizzera*; sentenza del 30 gennaio 2025, ricorsi nn. 51567/14, 39742/14, 74208/14 e 21215/15, *Cannavacciuolo e altri c. Italia*). Inoltre, il 29 settembre 2021 l'Assemblea parlamentare del Consiglio d'Europa ha presentato una proposta di protocollo addizionale alla CEDU sul diritto ad un ambiente salubre, pulito, sicuro e sostenibile, che consta di 11 articoli suddivisi in 5 sezioni.

graduali e differenziate a seconda della gravità e del grado di probabilità del verificarsi di impatti negativi su diritti umani, democrazia e Stato di diritto.

Tale approccio impone, dunque, una valutazione costante dei rischi e dei possibili effetti negativi e, a tal uopo, il trattato adotta un metodo proattivo, richiedendo agli Stati di prevenire i possibili abusi tecnologici, piuttosto che intervenire soltanto *ex post*, a violazione avvenuta.

Sul punto la Convenzione prevede un certo margine di flessibilità sulle modalità attuative delle disposizioni convenzionali, in considerazione, altresì, della diversità di sistemi giuridici, tradizioni e pratiche tra le Parti e dell'estrema varietà dei contesti in cui si fa uso dei sistemi di intelligenza artificiale, sia nel settore pubblico che in quello privato³¹.

Di fatto, consentire indifferentemente l'“adozione” o il “mantenimento in vigore” di misure appropriate già esistenti prima dell'entrata in vigore della Convenzione induce a pensare che, ai fini dell'attuazione nazionale³², entrambi gli approcci possano ritenersi ugualmente sufficienti.

In ogni caso ricade sulle Parti decidere come bilanciare gli interessi concorrenti, rilevanti in ogni ambito, tenendo conto delle specificità del settore pubblico (ad esempio con riferimento a migrazione, controllo delle frontiere, asilo, ...), così come di quello privato, del quadro normativo interno e dell'agenda nazionale per l'intelligenza artificiale, per garantire, in ogni caso la protezione e la promozione dei diritti umani, della democrazia e dello Stato di diritto.

4. Ambito applicativo delle disposizioni convenzionali e limiti

Nel quadro di una tutela globale, i concetti chiave della Convenzione sono in linea con i contenuti del già menzionato *AI Act*, a partire dalla definizione stessa di intelligenza artificiale (art. 2 della Convenzione quadro)³³.

Quest'ultima, come noto, è stata elaborata dall'OCSE per indicare un sistema automatizzato che «per obiettivi espliciti o impliciti, deduce dall'input che riceve come generare output quali previsioni, contenuti, raccomandazioni o decisioni che possono influenzare ambienti fisici o virtuali» (così art. 3, n. 1, regolamento)³⁴. Ivi la scelta di

³¹ L. COTINO HUESO, *El Convenio sobre inteligencia artificial, derechos humanos, democracia y Estado de Derecho del Consejo de Europa*, in *Revista Administración & Ciudadanía*, EGAP, 2024, ove l'A. sottolinea che «*The Convention not only has a symbolic and meta-legal value, but is also a normative instrument, with the capacity for quasiconstitutional integration into the legal systems of the States Parties (...)*».

³² Circa l'Italia cfr. la Strategia italiana per l'Intelligenza artificiale 2024-2026, https://www.agid.gov.it/sites/agid/files/2024-07/Strategia_italiana_per_l_Intelligenza_artificiale_2024-2026.pdf e il disegno di legge italiano sull'intelligenza artificiale (“DDL AI”), approvato dal Parlamento il 20 marzo 2025 che riprende l'impostazione dell'*AI Act*.

³³ J. ZILLER, *The Council of Europe Framework Convention on Artificial Intelligence vs. the EU Regulation: two quite different legal instruments*, in *Rivista Interdisciplinare sul Diritto delle Amministrazioni Pubbliche*, 2024, p. 202 ss.

³⁴ La definizione contenuta nella Convenzione quadro è sostanzialmente identica a quella resa dall'OCSE nel 2019 e aggiornata l'8 novembre 2023. Vedi relazione esplicativa che accompagna la definizione

mutuare anche la terminologia depone positivamente nella direzione di un'armonizzazione e attuazione coesa dei diversi strumenti relativi all'intelligenza artificiale all'interno degli ordinamenti giuridici nazionali degli Stati-parte.

Ratione materiae la Convenzione trova applicazione, in particolare, rispetto all'utilizzo dei sistemi di IA nel settore pubblico, da parte di autorità pubbliche³⁵ o da soggetti privati che agiscono per loro conto³⁶, lasciando invece ai firmatari la scelta su come regolare in modo conforme alla Convenzione le attività svolte da privati. Ciascuna Parte specifica, infatti, in una dichiarazione presentata al Segretario generale del Consiglio d'Europa, al momento della firma o del deposito del proprio strumento di ratifica/accettazione/approvazione o adesione, come intende attuare tale obbligo rispetto alle attività svolte nell'ambito del ciclo di vita dei sistemi di intelligenza artificiale da attori privati, indicando misure appropriate per adempiervi³⁷, con il limite, in ogni caso, di non poter derogare o circoscrivere l'applicazione degli obblighi internazionali assunti per proteggere i diritti umani, la democrazia e lo Stato di diritto.

Invero ciò è stato oggetto di critiche sin da subito, lamentando il fatto che la Convenzione potesse non applicarsi alle aziende, essendo gli Stati liberi di scegliere se farlo o meno, consentendo loro, di fatto, di sviluppare e utilizzare l'IA secondo i propri interessi. Dubbi, questi, sollevati, altresì, dal Garante europeo della protezione dei dati nell'ambito dell'Unione europea.

Ad ogni modo come si legge nella relazione esplicativa (con riferimento all'art. 3, punto 30), per le Parti che hanno scelto di non applicare i principi e gli obblighi della Convenzione quadro in relazione alle attività di altri attori privati, i redattori si aspettano che le posizioni di tali Parti mutino nel tempo man mano che i loro approcci alla regolamentazione del settore privato si evolvano.

aggiornata di un sistema di intelligenza artificiale nella Raccomandazione dell'OCSE sull'intelligenza artificiale (OCSE/LEGAL/0449, 2019, modificata nel 2023): *Explanatory memorandum on the updated OECD definition of an AI system, in OECD artificial intelligence papers, 2024*, n. 3. Tale definizione riflette un'ampia comprensione di cosa siano i sistemi di intelligenza artificiale, in particolare rispetto ad altri tipi di sistemi software tradizionali più semplici basati su regole definite esclusivamente da persone fisiche per eseguire automaticamente le operazioni. È intesa a garantire precisione e certezza giuridica, pur rimanendo sufficientemente astratta e flessibile tale da rimanere valida nonostante i futuri sviluppi tecnologici. Sebbene questa definizione fornisca un significato comune tra le Parti su cosa siano i sistemi di intelligenza artificiale, le Parti possono specificarla ulteriormente nei loro sistemi giuridici nazionali per ulteriore certezza e precisione giuridica, senza limitarne la portata. Per esempio, la versione della definizione di IA nel 2019 definiva come sistemi IA solo quelli che potevano fornire un output sulla base di una serie di obiettivi stabiliti da esseri umani («for a given set of human-defined objectives») dunque tenendo al di fuori di essa, per esempio, i sistemi di IA generati da altri sistemi IA.

³⁵ In linea con la Raccomandazione n. R (84) 15 del Comitato dei Ministri agli Stati membri relativa alla responsabilità civile del 18 settembre 1984, l'interpretazione condivisa dai redattori è che il termine "autorità pubblica" significa qualsiasi entità di diritto pubblico di qualsiasi tipo o livello (inclusi enti pubblici sovranazionali, statali, regionali, provinciali, municipali ed indipendenti) e qualsiasi persona privata quando esercita prerogative di autorità ufficiale.

³⁶ Attività per le quali le autorità pubbliche delegano le proprie responsabilità ad attori privati o dispongono loro di agire, come le attività di attori privati che operano ai sensi di un contratto con un'autorità pubblica o altra fornitura privata di servizi pubblici, nonché appalti e contratti pubblici.

³⁷ Poiché è importante per le Parti della Convenzione quadro sapere quali dichiarazioni sono state formulate, il Segretario generale del Consiglio d'Europa condividerà immediatamente le dichiarazioni ricevute con le altre Parti. Le Parti possono, in qualsiasi momento e nello stesso modo, modificare le proprie dichiarazioni.

Sono, invece, escluse dall'ambito applicativo le attività che rientrano nel ciclo di vita dei sistemi di IA connesse alla protezione degli interessi di sicurezza nazionale³⁸. Aspetto anch'esso criticabile, atteso che la Convenzione potrebbe non applicarsi qualora gli Stati consentano che l'IA venga sviluppata e utilizzata sotto le mentite spoglie della "sicurezza nazionale". Si pensi, a titolo esemplificativo, al riconoscimento facciale negli spazi pubblici, alla protezione delle frontiere tramite IA o alla scansione dei profili dei social media.

Tuttavia resta fermo che tali attività siano condotte in modo coerente con il diritto internazionale applicabile, compresi gli obblighi di diritto internazionale in materia di diritti umani, e nel rispetto delle proprie istituzioni e dei processi democratici; la sicurezza nazionale è, infatti, inclusa nell'ambito di diverse norme internazionali³⁹, le quali impongono che le misure volte a proteggerla devono essere previste dalla legge, rispettare l'essenza dei diritti umani e costituire una misura necessaria e proporzionata in una società democratica.

Inoltre, la Convenzione non si applica alle attività di ricerca e sviluppo relative a sistemi di intelligenza artificiale⁴⁰ non ancora resi disponibili per l'uso e che potenzialmente non interferiscano con diritti umani, democrazia e Stato di diritto, fermo restando che quest'ultime devono in ogni caso essere svolte in conformità con i diritti umani applicabili e il diritto interno, nonché con gli standard etici e professionali riconosciuti per la ricerca scientifica (art. 3, par. 3).

Infine, non rientrano nel campo di applicazione della Convenzione le questioni relative alla difesa nazionale (art. 3, par. 4), il che ricorda l'art. 1, lett. d, dello Statuto del Consiglio d'Europa⁴¹ ove si afferma che «*Le questioni attinenti alla Difesa Nazionale sono escluse dalla competenza del Consiglio d'Europa*».

Con riguardo ai limiti applicativi giova evidenziare che, nell'ambito dell'Unione europea, invece, il citato regolamento 2024/1689, che costituirà il principale atto giuridico dell'Unione ai fini dell'attuazione della Convenzione⁴², esclude dal suo ambito di

³⁸ Ciò mantiene nell'ambito della Convenzione quadro le attività relative ai sistemi di intelligenza artificiale "a duplice uso" nella misura in cui sono destinati a essere utilizzati per altri scopi non correlati alla protezione degli interessi di sicurezza nazionale delle Parti (così relazione esplicativa par. 32).

³⁹ Vedi la Convenzione del Consiglio d'Europa per la salvaguardia dei diritti dell'uomo e delle libertà fondamentali, la Convenzione americana sui diritti dell'uomo (Patto di San José), il Patto internazionale delle Nazioni Unite (ONU) sui diritti civili e politici e il Patto internazionale delle Nazioni Unite sui diritti economici, sociali e culturali.

⁴⁰ Così anche art. 2, par. 6 e par. 8 *AI Act*.

⁴¹ Statuto del Consiglio d'Europa, Londra, 5.V.1949.

⁴² Altre disposizioni pertinenti dell'*acquis* dell'Unione rilevanti ai fini dell'attuazione della Convenzione comprendono atti giuridici volti ad attuare i diritti fondamentali sanciti dalla Carta dei diritti fondamentali dell'Unione europea, quali la legislazione dell'Unione in materia di non discriminazione, comprese la direttiva 2000/43/CE del Consiglio, del 29 giugno 2000, che attua il principio della parità di trattamento fra le persone indipendentemente dalla razza e dall'origine etnica, in GU L 180 del 19.7.2000, p. 22; la direttiva 2000/78/CE del Consiglio, del 27 novembre 2000, che stabilisce un quadro generale per la parità di trattamento in materia di occupazione e di condizioni di lavoro, in GU L 303 del 2.12.2000, p. 16; l'*acquis* dell'Unione in materia di protezione dei dati, compresi il Regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio, del 27 aprile 2016, relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati e che abroga la direttiva 95/46/CE

applicazione i sistemi di IA immessi sul mercato, messi in servizio o utilizzati, con o senza modifiche, per scopi militari, di difesa o di “sicurezza nazionale” (art. 2) – la quale, come precisato anzitutto all’art. 4, par. 2 del Trattato sull’Unione europea (TUE), resta di competenza esclusiva di ciascuno Stato membro – nonché l’output dei sistemi di IA utilizzato nell’Unione solo per tali scopi, indipendentemente dal tipo di entità che svolge tali attività.

5. Obblighi generali e principi a garanzia di sistemi di intelligenza artificiale nel rispetto di diritti umani, democrazia e *rule of law*

In tale contesto la Convenzione prevede degli “obblighi generali” (artt. 4-5) per garantire, anzitutto, che i sistemi di intelligenza artificiale siano conformi agli obblighi di tutela dei diritti umani sanciti dal diritto internazionale⁴³ e nazionale. Invita, dunque, le

(Regolamento generale sulla protezione dei dati), in GU L 119 del 4.5.2016, p. 1) e il 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), in GU L 277 del 27.10.2022, p. 1, che mira ad assicurare un ambiente online sicuro, prevedibile e affidabile in cui siano rispettati i diritti fondamentali, tra cui il diritto alla libertà di espressione e il diritto di ricevere e diffondere informazioni; il Regolamento (UE) 2024/900 del Parlamento europeo e del Consiglio, del 13 marzo 2024, relativo alla trasparenza e al targeting della pubblicità politica, in GU L, 2024/900, 20.3.2024, ELI: <http://data.europa.eu/eli/reg/2024/900/oj>), e la legislazione sulla sicurezza dei prodotti e la legislazione sulla responsabilità per danno da prodotti difettosi, compresa la direttiva 85/374/CEE del Consiglio Direttiva 85/374/CEE del Consiglio del 25 luglio 1985 relativa al ravvicinamento delle disposizioni legislative, regolamentari ed amministrative degli Stati membri in materia di responsabilità per danno da prodotti difettosi, in GU L 210 del 7.8.1985, p. 29).

⁴³ Tra i principali strumenti e trattati internazionali sui diritti umani a livello globale: la Dichiarazione universale dei diritti umani del 1948, la Convenzione delle Nazioni Unite del 1965 sull’eliminazione di tutte le forme di discriminazione razziale, il Patto delle Nazioni Unite sui diritti civili e politici del 1966, il Patto delle Nazioni Unite sui diritti economici, sociali e culturali del 1966, la Convenzione delle Nazioni Unite del 1979 sull’eliminazione di tutte le forme di discriminazione contro le donne (CEDAW), la Convenzione delle Nazioni Unite del 1984 contro la tortura e altre pene o trattamenti crudeli, inumani o degradanti, la Convenzione delle Nazioni Unite sui diritti dell’infanzia del 1989, la Convenzione per la protezione di tutte le persone dalla sparizione forzata del 2006, la Convenzione delle Nazioni Unite sui diritti delle persone con disabilità del 2006. A livello regionale vedi: la Convenzione europea per la salvaguardia dei diritti dell’uomo e delle libertà fondamentali, la Carta sociale europea del 1961 e Carta sociale europea riveduta del 1966, la Convenzione del 1981 sulla protezione delle persone rispetto al trattamento automatizzato di dati a carattere personale, come modificata, la Convenzione europea del 1987 per la prevenzione della tortura e delle pene o trattamenti inumani o degradanti, la Convenzione del 1997 per la protezione dei diritti dell’uomo e della dignità dell’essere umano rispetto all’applicazione della biologia e della medicina: Convenzione sui diritti dell’uomo e la biomedicina (Convenzione di Oviedo), la Convenzione quadro del 1998 per la protezione delle minoranze nazionali, la Carta dei diritti fondamentali dell’Unione europea del 2000, la Convenzione del Consiglio d’Europa del 2005 sulla lotta contro la tratta di esseri umani, la Convenzione del Consiglio d’Europa del 2007 sulla protezione dei bambini contro lo sfruttamento e gli abusi sessuali (Convenzione di Lanzarote), la Convenzione del Consiglio d’Europa del 2011 sulla prevenzione e la lotta contro la violenza nei confronti delle donne e la violenza domestica (Convenzione di Istanbul). Vengono, in rilievo, altresì, altri strumenti digitali come: la Convenzione americana sui diritti dell’uomo del 1969 (Patto di San José), la Convenzione interamericana del 1985 per prevenire e punire la tortura, la Convenzione interamericana del 1994 sulla sparizione forzata delle persone, la Convenzione interamericana del 1994 sulla prevenzione, la punizione e l’eliminazione della violenza contro le donne, la Convenzione interamericana del 1999 sull’eliminazione di tutte le forme di

Parti a fare il punto sui propri obblighi, sulle fonti e sui meccanismi esistenti nel proprio ordinamento giuridico a tutela dei diritti umani, in linea, tra l'altro, con l'approccio descritto nell'art. 1, par. 2, di garantire che i quadri, le norme e i meccanismi esistenti continuino a proteggere e promuovere i diritti dell'uomo, in conformità agli obblighi internazionali vigenti, e siano sufficienti ed efficaci in risposta all'evoluzione dell'intelligenza artificiale.

Si tratta di obblighi di risultato⁴⁴ e non già di mezzi, rimettendo alle Parti la scelta delle modalità attuative, a partire dalla decisione, come sopra precisato, di adottare o di mantenere misure pertinenti.

In secondo luogo, nel rispetto della discrezionalità circa il *modus operandi*, la Convenzione chiede che i sistemi di intelligenza artificiale non siano utilizzati, altresì, per compromettere l'integrità, l'indipendenza e l'efficacia delle istituzioni, compreso il principio della separazione dei poteri, il rispetto dell'indipendenza della magistratura e l'accesso alla giustizia, introducendo, ad esempio, *bias* algoritmici che possano compromettere l'imparzialità decisionale (art. 5, par. 1); inoltre, impone di proteggere i processi democratici, tra cui l'equo accesso degli individui al dibattito pubblico e la loro partecipazione allo stesso, nonché la capacità di maturare liberamente opinioni (art. 5, par. 2). I redattori hanno così voluto indicare specifici contesti sensibili (il paragrafo 1 copre principalmente gli aspetti istituzionali rilevanti e il paragrafo 2 i processi democratici) in cui il potenziale utilizzo dell'intelligenza artificiale dovrebbe essere preceduto da un'attenta considerazione dei rischi per la democrazia e lo Stato di diritto⁴⁵ e accompagnato da norme e garanzie appropriate⁴⁶.

discriminazione contro le persone con disabilità, la Convenzione interamericana del 2013 contro il razzismo, la discriminazione razziale e le forme correlate di intolleranza, la Convenzione interamericana del 2015 sulla protezione dei diritti umani degli anziani.

⁴⁴ *Explanatory Report*, cit. par. 38.

⁴⁵ Nell'attuazione dei propri obblighi di tutela delle istituzioni e dei processi democratici ai sensi dell'articolo 5, le Parti potrebbero voler concentrarsi, ad esempio, sui rischi dei sistemi di intelligenza artificiale per: a) il principio della separazione dei poteri (esecutivo, legislativo e giudiziario); b) un sistema efficace di controlli ed equilibri tra i tre rami del governo, d) pluralismo politico (garantito in larga parte dalla tutela dei diritti umani, il cui rispetto è essenziale per una democrazia fiorente, come la libertà di espressione, la libertà di associazione e la libertà di riunione pacifica; e l'esistenza di media pluralisti e indipendenti e di una serie di partiti politici che rappresentano interessi e opinioni diversi) e un equo accesso e partecipazione al dibattito pubblico; g) il rispetto dello stato di diritto (che generalmente comprende i principi di legalità, certezza del diritto e non arbitrarietà) e il principio di accesso alla giustizia e alla sua corretta amministrazione; e h) rispetto del principio di indipendenza della magistratura.

⁴⁶ Raccomandazione CM/Rec(2024)5 del Comitato dei ministri del Consiglio d'Europa agli Stati membri in relazione agli aspetti etici e organizzativi sull'uso dell'IA e delle tecnologie digitali da parte dei Servizi penitenziari e di libertà vigilata. Il 9 ottobre 2024 il Consiglio d'Europa ha espresso ai suoi 46 Stati membri alcune raccomandazioni volte a garantire un uso dell'intelligenza artificiale e delle tecnologie digitali, da parte dei servizi penitenziari e di libertà vigilata, rispettoso dei diritti umani e della dignità dei detenuti e delle persone in libertà vigilata. In particolare, l'impiego dell'IA e delle tecnologie digitali deve essere rivolto alla riabilitazione dei detenuti, in ogni caso complementare e strumentale per migliorare anche la risocializzazione e ridurre la recidiva, senza mai sostituire la relazione e la decisione umana. Nel richiamare la citata Convenzione quadro del Consiglio d'Europa sull'Intelligenza artificiale, sui Diritti umani, sulla Democrazia e sullo Stato di diritto (CETS n. 225), agli Stati membri è raccomandato di tutelare tra l'altro la riservatezza e l'integrità dei dati personali e di disporre delle risorse necessarie per controllare e monitorare assiduamente tale uso.

Gli articoli da 6 a 13 dettano, quindi, una serie di principi generali comuni che le Parti aderenti si impegnano a rispettare relativamente alle attività del ciclo vita dei sistemi di IA (art. 6)⁴⁷, quali, *in primis* la dignità umana e l'autonomia individuale⁴⁸, per far sì che gli individui abbiano il controllo sull'uso e l'impatto delle tecnologie di intelligenza artificiale nelle loro vite e che la propria capacità di azione non ne risulti diminuita (art. 7).

Seguono la trasparenza e il controllo per i contenuti generati dall'AI specie per distinguerli da quelli generati dall'uomo, il cd. *oversight* (art. 8)⁴⁹. La trasparenza, in particolare, è di indubbio rilievo perché funzionale a garantire alle "persone interessate" di accedere a rimedi giuridici efficaci in caso di violazioni correlate all'uso dell'IA (art. 14). A tal proposito la Convenzione chiede agli Stati firmatari di garantire che tali sistemi operino secondo standard di trasparenza tali da permettere agli utenti di avere contezza del processo decisionale algoritmico. E ciò va ben oltre il concetto di mera accessibilità delle informazioni, implicando una vera e propria comprensione dell'*iter* decisionale automatizzato, sì da consentire agli interessati di poter contestare decisioni che incidono

⁴⁷ L'art. 6 chiarisce che i principi dovrebbero essere incorporati negli approcci nazionali delle Parti alla regolamentazione dei sistemi di intelligenza artificiale. In quanto tali, sono volutamente redatti a un elevato livello di generalità, con l'intenzione che siano requisiti sovraordinati che possono essere applicati in modo flessibile in una varietà di contesti in rapido cambiamento, tenendo conto, tra l'altro, che in virtù dei rispettivi obblighi internazionali in materia di diritti umani ciascuna Parte dispone già di un regime giuridico dettagliato di protezione dei diritti umani con il proprio insieme di regole, principi e pratiche riguardanti l'ambito, il contenuto dei diritti e le possibili restrizioni, deroghe o eccezioni a tali diritti, nonché il funzionamento dei meccanismi di supervisione e di applicazione applicabili.

⁴⁸ Come si legge nella relazione esplicativa, l'autonomia individuale è un aspetto importante della dignità umana e si riferisce alla capacità degli individui di autodeterminarsi, ovvero alla loro capacità di fare scelte e prendere decisioni, anche senza coercizione, e di vivere la propria vita liberamente. Nel contesto dell'intelligenza artificiale, l'autonomia individuale richiede che gli individui abbiano il controllo sull'uso e l'impatto delle tecnologie di intelligenza artificiale nelle loro vite e che la loro capacità di azione e autonomia non ne risultino diminuite. La regolamentazione incentrata sull'uomo riconosce l'importanza di consentire agli individui di modellare le proprie esperienze con l'intelligenza artificiale, assicurando che queste tecnologie migliorino piuttosto che violare la loro autonomia.

⁴⁹ La disposizione prevede anche misure relative all'identificazione dei contenuti generati dall'IA al fine di evitare il rischio di inganno e consentire la distinzione tra contenuti autentici generati dall'uomo e contenuti generati dall'IA, poiché diventa sempre più difficile per le persone identificarli. Il principio di trasparenza di cui all'articolo 8 si riferisce all'apertura e alla chiarezza nella governance delle attività all'interno del ciclo di vita dei sistemi di intelligenza artificiale e significa che i processi decisionali e il funzionamento generale dei sistemi di intelligenza artificiale dovrebbero essere comprensibili e accessibili agli attori appropriati dell'intelligenza artificiale e, ove necessario e appropriato, alle parti interessate pertinenti. In alcuni casi, ciò potrebbe anche riferirsi alla fornitura di informazioni aggiuntive, tra cui, ad esempio, sugli algoritmi utilizzati, soggetti a sicurezza, proprietà commerciale e intellettuale e altre considerazioni, come dettagliato nel paragrafo 62 di seguito. I mezzi per garantire la trasparenza dipenderebbero da molti fattori diversi come, ad esempio, il tipo di sistema di intelligenza artificiale, il contesto del suo utilizzo o il suo ruolo e il *background* dell'attore pertinente o della parte interessata. Inoltre, le misure pertinenti includono, ove appropriato, la registrazione di considerazioni chiave come la provenienza dei dati, le metodologie di formazione, la validità delle fonti di dati, la documentazione e la trasparenza sui dati di formazione, test e convalida utilizzati, gli sforzi di mitigazione del rischio e i processi e le decisioni implementati, al fine di favorire una comprensione completa di come gli *output* del sistema di intelligenza artificiale vengono derivati e hanno un impatto sui diritti umani, sulla democrazia e sullo stato di diritto. Ciò contribuirà in particolare a garantire la responsabilità e consentirà alle persone interessate di contestare l'uso o i risultati del sistema di intelligenza artificiale, ove e come applicabile.

su diritti e/o libertà⁵⁰. Si pensi, ad esempio, ai pazienti che si sottopongono a prestazioni mediche rese con l'ausilio dell'IA, i quali, secondo il citato articolo 14, vanterebbero un "diritto alla spiegazione", in una prospettiva rimediabile.

Anche nell'*AI Act* si rinviene un diritto alla spiegazione con riguardo a sistemi di IA ad alto rischio⁵¹, il che va a rafforzare quanto già previsto all'art. 22 del regolamento generale sulla protezione dei dati (GDPR)⁵², rubricato *Processo decisionale automatizzato relativo alle persone fisiche, compresa la profilazione*.

Procedendo, l'art. 9 in tema di *accountability* e *responsibility* impone agli Stati l'adozione di misure volte a garantire l'individuazione delle rispettive responsabilità nell'implementazione dei sistemi di IA; in special modo si tratta di comprendere come il sistema abbia generato determinate decisioni, specie in situazioni in cui gli algoritmi possano incidere su diritti umani, democrazia e *rule of law*, assicurando l'identificazione del soggetto, individuale o collettivo, chiamato a rispondere degli effetti lesivi.

E ancora, la Convenzione pone l'accento sull'impegno di salvaguardare: la non discriminazione e l'uguaglianza compresa quella di genere⁵³, ai sensi del diritto internazionale⁵⁴ e nazionale applicabile, per evitare che la progettazione del ciclo di vita

⁵⁰ La relazione esplicativa chiarisce due importanti accezioni concettuali del principio di trasparenza, declinabile in termini di spiegabilità e interpretabilità. In particolare, la spiegabilità viene intesa come «capacità di fornire, subordinatamente alla fattibilità tecnica e tenendo conto dello stato dell'arte generalmente riconosciuto, spiegazioni sufficientemente comprensibili sul motivo per cui un sistema di intelligenza artificiale fornisce informazioni, produce previsioni, contenuti, raccomandazioni o decisioni». Il canone di interpretabilità, invece, «si riferisce alla capacità di comprendere come un sistema di intelligenza artificiale effettua le sue previsioni o decisioni o, in altre parole, la misura in cui i risultati dei sistemi di intelligenza artificiale possono essere resi accessibili e comprensibili esperti e non esperti. Si tratta di rendere il funzionamento interno, la logica e i processi decisionali dei sistemi di intelligenza artificiale comprensibili e accessibili agli utenti umani, inclusi sviluppatori, parti interessate, utenti finali e persone interessate».

⁵¹ Art. 68 *AI Act*, considerando n. 93 e n. 171.

⁵² Regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio, del 27 aprile 2016, relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati e che abroga la direttiva 95/46/CE (regolamento generale sulla protezione dei dati), in GU L 119 del 4 maggio 2016, pp. 1-88.

⁵³ Per approfondimenti cfr. P. ZUDDAS, *Intelligenza artificiale e discriminazioni*, in AA.VV. (a cura di), *Liber amicorum per Pasquale Costanzo – Diritto Costituzionale in trasformazione, vol. I – Costituzionalismo, Reti e Intelligenza artificiale*, Genova, 2020, p. 457 ss.; M. D'AMICO, C. NARDOCCI, *Discriminazioni, Donne e Intelligenza Artificiale*, in G. CERRINA FERONI, C. FONTANA, E.C. RAFFIOTTA (a cura di), *AI Anthology*, Bologna, 2022.

⁵⁴ A livello globale, gli articoli 2, 24 e 26 dell'ICCPR; articoli 2, 3 e 7 del Patto internazionale sui diritti dell'uomo e delle libertà fondamentali; strumenti giuridici specializzati quali l'ICERD, la CEDAW, la UNCRC e la UNCRPD; a livello regionale, l'articolo 14 della CDU e il suo Protocollo n. 12; paragrafi 20 e 27 della Parte I, l'articolo 20 della Parte II e l'articolo E della Parte V della CSE; Strumenti giuridici specializzati del Consiglio d'Europa, come l'articolo 4 della Convenzione quadro Convenzione per la protezione delle minoranze nazionali e articolo 4 della Convenzione di Istanbul; Titolo III della Carta dei diritti fondamentali dell'Unione europea, i Trattati dell'UE (ad esempio, l'articolo 2 del Trattato sull'Unione europea, l'articolo 10 del Trattato sul funzionamento dell'Unione europea), il diritto derivato dell'UE e la pertinente giurisprudenza della Corte di giustizia dell'Unione europea; Articolo 24 del Patto di San José; e Strumenti giuridici specializzati, quali la Convenzione interamericana del 1999 sull'eliminazione di tutte le forme di discriminazione contro le persone con disabilità; la Convenzione interamericana del 2013 contro il razzismo, la discriminazione razziale e le forme di intolleranza connesse e la Convenzione interamericana del 2015 sulla protezione dei diritti umani delle persone anziane.

dei sistemi di IA possa provocare pregiudizi “incorporati intenzionalmente o inavvertitamente” nell’addestramento dei relativi algoritmi, generando distorsioni applicative errate, come disuguaglianze o lesioni dei diritti delle minoranze (art. 10); il diritto alla *privacy* e la protezione dei dati personali (art. 11), in ossequio, altresì, a quanto previsto dall’art. 8 della Convenzione europea⁵⁵; l’affidabilità dei sistemi di intelligenza artificiale (cd. *reliability*) e la fiducia negli output da essi prodotti (art. 12); e, come anticipato, l’innovazione sicura tramite sperimentazione, prevedendo l’istituzione di ambienti controllati per lo sviluppo, la sperimentazione e la prova di sistemi di intelligenza artificiale sotto la supervisione di autorità competenti (art. 13).

Tanto esposto, gli obblighi generali e i principi sopraelencati hanno il pregio di arginare difetti e limiti di tecniche palesemente all’avanguardia ma al contempo, come visto, potenzialmente insite di contraddizioni.

6. Rimedi e gestione dei rischi e degli impatti negativi

A seguire, il Capo IV, nell’ottica di una tutela effettiva, reca una disposizione sui rimedi per le violazioni dei diritti umani (art. 14).

Sul punto la Convenzione quadro, muovendo dalla constatazione che ciascuno Stato-parte è già vincolato al rispetto di obblighi internazionali e nazionali in relazione ai diritti umani, alla democrazia e allo Stato di diritto, chiede alle Parti di garantire la disponibilità dei rimedi esistenti anche nel caso di violazioni derivanti dall’uso di sistemi di IA⁵⁶, in un’ottica di continuità tra violazioni imputabili all’“intelligenza naturale/umana”, e violazioni imputabili all’intelligenza artificiale.

Si tratta, pertanto, di “obblighi” di fornire tutela a carico degli Stati-parte e non già di un sistema di rimedi a livello del Consiglio d’Europa. A tal uopo si impone loro di adottare o mantenere misure volte a garantire: le informazioni pertinenti e necessarie relative ai sistemi di intelligenza artificiale aventi il potenziale di “incidere in modo significativo sui diritti umani”, nonché l’effettiva possibilità per le persone interessate di presentare un reclamo alle autorità competenti, nel rispetto delle garanzie procedurali (art. 15). Ciò risponde all’esigenza di compensare quell’asimmetria informativa connaturata ai sistemi di IA per la loro complessità tecnica, ovvero quel significativo squilibrio nell’accesso, nella comprensione o nel controllo delle informazioni tra le diverse parti coinvolte, al fine di rendere i rimedi esistenti, efficaci e accessibili, in conformità, altresì,

⁵⁵ In tema di *privacy*, a livello globale v. l’art. 17 dell’ICCPR e a livello regionale l’art. 8 della CEDU, l’articolo 8 della Carta dell’UE e l’art. 11 del Patto di San José. L’art. 11, sottoparagrafo (b) fa anche esplicito riferimento ad altre “garanzie e tutele” di cui gli individui solitamente godono in virtù di tali leggi, standard e quadri normativi. Uno di questi strumenti è la Convenzione 108+ del Consiglio d’Europa, che copre sia il settore pubblico che quello privato ed è aperta all’adesione degli Stati a livello globale. A livello dell’UE, il Regolamento generale sulla protezione dei dati (Regolamento (UE) 2016/679, “GDPR”).

⁵⁶ Per sottolineare il collegamento e garantire la complementarità con i meccanismi di protezione dei diritti umani internazionali e nazionali applicabili, la disposizione utilizza la terminologia giuridica “violazioni dei diritti umani” a cui si fa riferimento nell’art. 2 dell’ICCPR, negli artt. 13, 34, 41 e 46 della CEDU e negli artt. 25 e 63 del Patto di San José.

ai già citati principi di trasparenza e controllo di cui all'art. 8 e al principio di *accountability* e responsabilità *ex art.* 9.

E qui giova ricordare la possibilità di eccezioni, limitazioni o deroghe agli obblighi di trasparenza a tutela dell'ordine pubblico, della sicurezza e di altri rilevanti interessi pubblici come previsto, ove necessario, da strumenti internazionali applicabili.

Inoltre, come sottolineato nell'*Explanatory Report*, le espressioni “incidere in modo significativo sui diritti umani” nel sottoparagrafo a) del par. 2 dell'art. 14 e “incidere in modo significativo sul godimento dei diritti umani” nel par. 1 dell'art. 15 introducono entrambe un requisito soglia⁵⁷, il che significa che i citati articoli 14 e 15 non si applicano automaticamente a tutti i sistemi di intelligenza artificiale rientranti nell'ambito di applicazione dell'art. 3 della Convenzione, bensì solo a quelli aventi un effetto o impatto significativo sui diritti umani.

Dal canto suo, l'art. 15 prevede che gli Stati-parte assicurino che le garanzie procedurali applicabili sempre in materia di diritti umani, sia a livello internazionale che nazionale, operino e siano efficaci anche nel contesto dell'intelligenza artificiale. Ad esempio, esse dovrebbero includere la supervisione umana da parte di operatori competenti, compresa la revisione *ex ante* o *ex post* della decisione adottata.

Infine, per evitare il rischio di manipolazione e inganno, l'art. 15 prevede espressamente che ciascuna Parte si impegni a garantire che, ove opportuno in base al contesto, le persone che si interfacciano con sistemi di intelligenza artificiale siano messe al corrente che stiano interagendo con tali sistemi e non con un essere umano.

All'interno del testo sono presenti, altresì, disposizioni su previsioni/misure per l'identificazione, valutazione, prevenzione e mitigazione dei rischi posti da tutti i sistemi di IA, senza distinzioni, tenendo conto degli impatti reali e potenziali su diritti umani, democrazia e Stato di diritto (art. 16), a differenza di quanto avviene nell'ambito del diritto dell'Unione europea, ove la gestione del rischio è in via generale e poi ulteriormente specificata con riguardo alla valutazione d'impatto sui diritti fondamentali circa i sistemi c.d. “ad alto rischio”⁵⁸.

In ogni caso trattasi di misure graduate e differenziate che tengono in debita considerazione il contesto e l'uso previsto dei sistemi di intelligenza artificiale, in particolare per quanto concerne i rischi per i diritti umani, la democrazia e lo Stato di diritto, il che consente flessibilità alle Parti negli approcci e nelle metodologie da seguire. Ad esempio, le Parti possono scegliere di agire a livello normativo prescrivendo diverse categorie di classificazione del rischio e/o a livello operativo incaricando soggetti responsabili per le attività all'interno del ciclo di vita dei sistemi di intelligenza artificiale, oppure concentrarsi solo su determinate categorie predefinite di sistemi di IA.

Sul punto il Consiglio d'Europa si astiene dal fornire una nozione di rischio, a differenza del legislatore dell'UE⁵⁹, che distingue per di più i sistemi di AI in base al loro

⁵⁷ *Explanatory Report*, cit., par. 101.

⁵⁸ Art. 27 *AI Act*.

⁵⁹ Art. 3, n. 2, *AI Act*: “rischio”: la combinazione della probabilità del verificarsi di un danno e la gravità del danno stesso”.

livello di rischio: inaccettabile⁶⁰, alto⁶¹, specifico per la trasparenza⁶² e minimo⁶³ e, pone obblighi più stringenti qualora si tratti di modelli classificati con rischio sistemico (art. 51 ss.).

Tali misure, come anticipato, devono tenere debitamente conto della gravità e della probabilità dei potenziali impatti, di coloro i cui diritti potrebbero essere lesi e, includere il monitoraggio dei rischi e degli effetti negativi su diritti umani, democrazia e Stato di diritto, così come la relativa documentazione. In particolare, esse dovrebbero richiedere, ove opportuno, la sperimentazione dei sistemi di intelligenza artificiale prima di renderli disponibili per il primo utilizzo o qualora subiscano modifiche significative, con la possibilità per gli Stati, in linea con il regolamento UE, di vietare determinate tipologie di intelligenza artificiale che costituiscano, sulla base degli strumenti di gestione del rischio, un pericolo per la democrazia, i diritti fondamentali e lo Stato di diritto.

Emerge, dunque, come, da un lato si voglia garantire un approccio uniforme all'identificazione, analisi e valutazione dei rischi e degli impatti di tali sistemi e, dall'altro, si sia consapevoli del fatto che le Parti siano nella posizione migliore per scelte normative pertinenti all'adempimento degli obblighi di cui alla Convenzione, – nell'ottica, del resto, di una protezione convenzionale sussidiaria – tenendo conto dei loro specifici contesti legali, politici, economici, sociali, culturali e tecnologici.

⁶⁰ Capo II, art. 5 *AI Act*. Sistemi di AI che pongono rischi inaccettabili, la cui commercializzazione e uso sono esplicitamente vietati nel mercato interno dell'UE. Tra questi rientrano i sistemi che: utilizzano tecniche manipolative nei confronti delle persone; sfruttano categorie di soggetti considerati vulnerabili; classificano le persone o i gruppi in base al comportamento sociale o alle caratteristiche personali in presenza di specifiche condizioni; valutano il rischio di condotte criminali; creano database di riconoscimento facciale attraverso attività di scraping; categorizzano le persone in base ai dati biometrici e realizzano l'identificazione biometrica in tempo reale (con alcune eccezioni).

⁶¹ Capo III, Sezione 1 *Classificazione dei sistemi di IA come «ad alto rischio»*, artt. 6 e 7 *AI Act*. I sistemi ad alto rischio potranno essere commercializzati, purché siano rispettati una serie di requisiti (Sezione 2 *Requisiti per i sistemi di IA ad alto rischio*), tra i quali rientrano l'adozione di un sistema di gestione dei rischi (art. 9), la redazione e l'aggiornamento della documentazione tecnica (art. 10), la predisposizione di misure tecniche che garantiscano, per l'utente, la possibilità di comprendere l'output del dispositivo (art. 13) e di supervisionarlo efficacemente (art. 14). Sistemi di IA ad alto rischio, sono, ad esempio, i software medici basati sull'IA o i sistemi di IA utilizzati per la selezione e l'assunzione di personale. Sul punto E. CIRONE, *L'AI Act e l'obiettivo (mancato?) Di promuovere uno standard globale per la tutela dei diritti fondamentali*, in *Quaderni AISDUE*, fasc. spec. 2024, n. 2, p. 64, ove l'A. evidenzia che ciò che viene particolarmente contestato è la probabilità che la valutazione del rischio per i diritti fondamentali segua gli stessi metodi di valutazione del rischio utilizzati per la sicurezza dei prodotti, aumentando così il pericolo di negligenza nella protezione dei diritti fondamentali. Infatti, a differenza del settore della sicurezza dei prodotti, dove il rischio può essere quantificato mediante calcoli computazionali, per i diritti fondamentali non è possibile riferirsi a mere valutazioni numeriche.

⁶² Capo IV, art. 50 *AI Act*. Qualora una persona fisica interagisca con questa tipologia di AI è stabilito l'obbligo di informare chi interessato della natura artificiale del sistema e dei contenuti generati dallo stesso. I sistemi come i chatbot devono informare chiaramente gli utenti che stanno interagendo con una macchina, mentre alcuni contenuti generati dall'IA devono essere etichettati come tali.

⁶³ Capo X. Questi sistemi non sono soggetti ad ulteriori obblighi oltre la normativa di settore e quanto già previsto a livello di regolamentazione europea. L'*AI Act* incoraggia comunque l'adozione di codici di condotta intesi a promuovere l'applicazione volontaria anche a questi sistemi di AI i requisiti previsti per i sistemi ad alto rischio (art. 95).

Resta fermo, in ogni caso, che l'attuazione delle disposizioni della Convenzione sia garantita senza alcuna discriminazione (art. 17)⁶⁴, tenendo debitamente conto delle esigenze e vulnerabilità specifiche in relazione al rispetto dei diritti delle persone con disabilità e dei bambini (art. 18)⁶⁵.

Infine, si vuole garantire che decisioni su questioni importanti relative ai sistemi di intelligenza artificiale siano, ove opportuno, debitamente prese a seguito di un dibattito pubblico e una consultazione multilaterale alla luce delle implicazioni sociali,

⁶⁴ Il significato di discriminazione nell'art. 17 è identico a quello stabilito nel diritto internazionale applicabile, come, tra gli altri, l'art. 26 dell'ICCPR, l'art. 2 dell'ICESCR, l'art. 14 della CEDU e il suo Protocollo n. 12, l'articolo 24 del Patto di San José e l'articolo E della ESC, se e come applicabile alle Parti della Convenzione quadro. Nel loro insieme, queste disposizioni coprono un'ampia gamma di motivi di non discriminazione che sono collegati alle caratteristiche personali degli individui, alle circostanze o all'appartenenza a un gruppo. Non tutti questi motivi sono esplicitamente dichiarati o formulati in modo identico nei trattati sui diritti umani da cui le Parti della presente Convenzione quadro possono essere vincolate. Tali trattati di solito contengono elenchi aperti di tali motivi, come interpretati dalla giurisprudenza di corti internazionali competenti come la Corte europea e la Corte interamericana dei diritti dell'uomo e nella prassi pertinente di organismi internazionali competenti, come il Comitato per i diritti umani delle Nazioni Unite. Possono, quindi, esserci delle variazioni tra i vari regimi internazionali sui diritti umani applicabili alle diverse Parti. Come per altre convenzioni e trattati sui diritti umani, anche qui l'approccio della Convenzione quadro non è quello di creare nuovi obblighi sui diritti umani o di ridurre, estendere o altrimenti modificare la portata o il contenuto degli obblighi internazionali in materia di diritti umani applicabili a una Parte. In tema di discriminazione e I.A. cfr. M. SAVIANO, C. BASSANO, *Parità di genere: rischi e opportunità nell'era dell'Intelligenza*, in A. DI STASI (a cura di), *Dalla non discriminazione alle pari opportunità*, cit., p. 211, ove si sottolinea che i rischi di discriminazione e di trattamento iniquo legati all'uso dell'IA possono non solo riprodurre le disuguaglianze esistenti ma addirittura di amplificarle. Tali rischi sono ricondotti a due cause principali. Da un lato, al fatto che le banche dati per l'apprendimento dell'IA, dato l'uso di insiemi di dati condivisi e riutilizzabili, tendono a riprodurre e a mantenere algoritmi inizialmente discriminatori. Dall'altro, al fatto che sviluppatori, progettisti e formatori di IA, proiettando, nel loro lavoro, la propria rappresentazione della realtà o della società, possono riflettere pregiudizi e stereotipi inconsci e causare comportamenti discriminatori da parte dei sistemi di IA che sviluppano. E, al riguardo, si sottolinea che, di recente, sono state intraprese diverse iniziative per la realizzazione di sistemi di IA equi, come lo sviluppo di database di apprendimento dell'Intelligenza Artificiale più diversificati e l'emanazione di principi e guide che comprendono l'etica e i principi per la diversità, l'equità e l'inclusione (Diversity, Equality, Inclusion - DEI) per lo sviluppo, la distribuzione, l'acquisizione, l'uso e la governance di IA eque e affidabili (Commissione europea, 2020).

⁶⁵ Questa disposizione stabilisce l'obbligo per le Parti, nel contesto delle attività nel ciclo di vita dei sistemi di intelligenza artificiale, di tenere in debita considerazione "le esigenze e le vulnerabilità specifiche in relazione al rispetto dei diritti delle persone con disabilità e dei bambini" e, a tale riguardo, è direttamente correlata alle disposizioni e al regime giuridico della Convenzione delle Nazioni Unite sui diritti delle persone con disabilità (United Nations Convention on the Rights of Persons with Disabilities - UNCRPD) e della Convenzione sui diritti del fanciullo (UN Convention on Rights of a Child – UNCRC), nonché alla legislazione nazionale applicabile di ciascuna Parte sui diritti delle persone con disabilità e sui diritti del bambino. È stato inserito un riferimento esplicito alla legislazione nazionale applicabile sui diritti del bambino e sui diritti delle persone con disabilità, in particolare, al fine di prendere in considerazione la situazione di qualsiasi Parte della Convenzione quadro che non abbia ratificato la UNCRC o la UNCRPD, ma che abbia comunque una legislazione nazionale che garantisca il godimento di tali diritti. Come si legge nella relazione esplicativa, in considerazione del grave rischio che le tecnologie di intelligenza artificiale possano essere utilizzate per facilitare lo sfruttamento sessuale e l'abuso sessuale dei minori, e dei rischi specifici che ciò comporta per i minori, nel contesto dell'attuazione di questa disposizione i redattori hanno preso in considerazione gli obblighi stabiliti dalla Convenzione di Lanzarote, dal Protocollo opzionale alla Convenzione delle Nazioni Unite sui diritti dell'infanzia sulla vendita di bambini, la prostituzione minorile e la pornografia minorile, e dall'Osservazione generale n. 25 alla Convenzione sui diritti dell'infanzia delle Nazioni Unite sui diritti dell'infanzia in relazione all'ambiente digitale.

economiche, legali, etiche, ambientali e di altro tipo, con esperti del settore, accademici, istituzioni nazionali, società civile (art. 19)⁶⁶.

A tal proposito si incoraggia un'adeguata alfabetizzazione e l'acquisizione di competenze digitali necessarie per tutti i segmenti della popolazione (art. 20), comprese le competenze specialistiche per i responsabili dell'identificazione, valutazione, prevenzione e attenuazione dei rischi posti dai sistemi di intelligenza artificiale (ad es. magistratura, autorità nazionali di vigilanza, autorità per la protezione dei dati, organismi per l'uguaglianza e i diritti umani, difensori civici, autorità per la tutela dei consumatori, fornitori e utenti di intelligenza artificiale).

Resta fermo che nessuna disposizione della Convenzione deve essere interpretata nel senso di limitare, derogare o incidere sui diritti umani (art. 21), né si esclude una tutela più ampia (art. 22) rispetto alle norme minime previste, facendo sì che la Convenzione medesima coesista armoniosamente con altri trattati e strumenti internazionali sui diritti umani.

7. A proposito del meccanismo di *follow-up* e della cooperazione

Degne di nota sono, in ultimo, le disposizioni che mirano a garantire l'effettiva attuazione della Convenzione da parte degli Stati attraverso un meccanismo di *follow-up*, oltre che di cooperazione, preannunciato all'art. 1, par. 3⁶⁷ e affidato ad un organismo *ad hoc*, quale la Conferenza delle parti (capitolo VII, da 23-35).

Quest'ultima garantirà la partecipazione paritaria di tutti gli Stati-parte al processo decisionale e alla procedura di *follow-up* della Convenzione e rafforzerà, altresì, la cooperazione tra le stesse al fine di garantire un'attuazione corretta ed efficace delle disposizioni convenzionali.

Tale organismo opererà alla stregua di un regolamento interno da adottarsi entro dodici mesi dall'entrata in vigore della Convenzione; sarà convocata dal Segretario Generale del Consiglio d'Europa ogniqualvolta necessario e, in ogni caso, quando la maggioranza delle Parti o il Comitato dei Ministri lo richiederà, all'insegna, dunque, della flessibilità, come emerge dal fatto che non è indicato alcun requisito temporale per la convocazione stessa.

Essa vanterà le tradizionali competenze di *follow up* e, pertanto, svolgerà un ruolo in merito all'attuazione effettiva della Convenzione quadro formulando proposte volte a

⁶⁶ L'art. 19 lascia alle Parti la determinazione degli argomenti, della frequenza e delle altre modalità di tali consultazioni alla luce delle implicazioni sociali, economiche, legali, etiche, ambientali e di altre implicazioni rilevanti. Ad esempio, gli Stati possono organizzare sondaggi e questionari, workshop pubblici, focus group, giurie di cittadini e sondaggi deliberativi, comitati di esperti e comitati consultivi, udienze pubbliche, conferenze nazionali e internazionali o combinazioni di quanto sopra. La valutazione finale e l'incorporazione dei risultati di tali discussioni e consultazioni nelle iniziative politiche pertinenti potrebbero anche essere adeguatamente e opportunamente comunicate alle parti interessate pertinenti.

⁶⁷ L'art. 1, par. 3: «Per garantire un'attuazione efficace della Convenzione quadro, la Convenzione quadro prevede l'istituzione di un meccanismo di follow-up e prevede la cooperazione internazionale».

facilitare o migliorare l'esecuzione, come pure proposte di emendamento conformemente all'art. 28, par. 1. Rivestirà, altresì, un ruolo consultivo rispetto alla Convenzione, con raccomandazioni specifiche su questioni relative alla sua interpretazione o applicazione, tra cui, ad esempio, suggerire interpretazioni di termini giuridici contenuti all'interno della Convenzione. Tali raccomandazioni, sebbene non saranno di natura giuridicamente vincolante, potranno ritenersi comunque espressione congiunta dell'opinione delle Parti su un dato argomento, da tenere in debita considerazione nell'applicazione della Convenzione medesima.

E ancora, la Conferenza fungerà da *forum* per facilitare lo scambio di informazioni su significative evoluzioni sul piano normativo, sociale, politico o tecnologico in relazione all'applicazione delle disposizioni della Convenzione e alle attività di cooperazione internazionale *ex art. 25*; agevolerà, ove necessario, la risoluzione amichevole delle controversie relative all'applicazione delle sue disposizioni, a titolo consultivo e non vincolante; faciliterà la cooperazione con le parti interessate, tra cui organizzazioni non governative e altri organismi; si avvarrà, altresì, della consulenza di esperti, in considerazione dell'oggetto altamente tecnico della Convenzione (art. 23, par. 6).

Vigerà, inoltre, un obbligo di segnalazione *ex art. 24* che imporrà a ciascuna Parte di fornire alla Conferenza delle parti, entro i primi due anni successivi alla sua adesione, e in seguito periodicamente, un rapporto contenente i dettagli delle attività intraprese per dare esecuzione alla Convenzione⁶⁸, al fine di consentire un aggiornamento regolare e, nell'ottica della cooperazione, condividere informazioni su misure e prassi intraprese.

E qui giova evidenziare la rilevanza attribuita nel testo convenzionale alla cooperazione (art. 25) al fine di perseguire lo scopo, ormai noto, di garantire che le attività nell'ambito del ciclo vita dei sistemi di intelligenza artificiale siano pienamente coerenti con i diritti umani, la democrazia e lo Stato di diritto⁶⁹; cooperazione, estesa anche agli Stati che non sono ancora Parti della Convenzione, offrendo loro supporto a che si conformino ai principi dettati, al fine di incoraggiarne l'adesione. Essa si sostanzia in special modo nello scambio di informazioni pertinenti e utili su aspetti relativi all'intelligenza artificiale che possano avere ricadute significative, positive o negative, sul godimento dei diritti umani, sul funzionamento della democrazia e sul rispetto dello Stato di diritto, compresi i rischi e gli effetti emersi nei contesti di ricerche sull'IA e relative al settore privato. Anche in questo caso le Parti sono incentivate a coinvolgere, nello scambio di informazioni, attori non statali (accademici, rappresentanti dell'industria e organizzazioni della società civile), oltre – come già sottolineato – gli Stati che non sono Parti.

⁶⁸ Cfr. art. 3, par. 1, lett. *a e b*.

⁶⁹ Report of the UN Secretary-General's High-level Panel on Digital Cooperation “*The age of digital interdependence*”, <https://www.un.org/en/pdfs/DigitalCooperation-report-for%20web.pdf>, ove si muove dalla consapevolezza che le opportunità di progresso umano nell'era digitale superino in ultima analisi le sfide se ci si unisce in uno spirito di cooperazione e inclusività.

E sempre nell'ottica di garantire un'effettiva attuazione della Convenzione, l'art. 26 impone alle Parti di adottare o mantenere organismi già esistenti per vigilare sul rispetto degli obblighi previsti, con imparzialità e indipendenza – sia riguardo al potere esecutivo che legislativo – disponendo di poteri, competenze e risorse necessarie per assolvere efficacemente i propri compiti.

Infine, nel caso di controversie relative all'applicazione o all'interpretazione della Convenzione si impone alle Parti l'obbligo di ricercare una risoluzione pacifica, potendosi rivolgere di comune accordo, come visto, anche alla Conferenza delle Parti.

8. La partecipazione attiva dell'Unione europea ai negoziati della Convenzione quadro: spunti di riflessione

Con particolare riguardo al coinvolgimento dell'Unione europea, come sopra anticipato, ai sensi della decisione (UE) 2022/2349 del Consiglio, del 21 novembre 2022, *che autorizza l'avvio di negoziati a nome dell'Unione europea per una convenzione del Consiglio d'Europa sull'intelligenza artificiale, i diritti umani, la democrazia e lo Stato di diritto*⁷⁰, il Consiglio a suo tempo ha autorizzato la Commissione ad avviare negoziati a nome dell'Unione, la quale ha così preso parte alla redazione del testo convenzionale, avvalendosi anche del contributo di rappresentanti dell'Agenzia dell'Unione europea per i diritti fondamentali e del Garante europeo della protezione dei dati.

Decisione questa, posta in essere in ambito UE, nel corso dei negoziati da parte del Parlamento europeo e del Consiglio relativi all'*AI Act* – poi adottato, come noto, il 13 giugno 2024⁷¹ ed entrato in vigore il 1° agosto 2024 – nella consapevolezza che l'ambito di applicazione della Convenzione coincidesse in larga misura con quello dell'allora proposta legislativa⁷²; entrambi gli atti mirano, infatti, – è ormai risaputo – a stabilire regole applicabili alla progettazione, allo sviluppo e all'applicazione di sistemi di IA forniti e utilizzati da soggetti pubblici o privati (considerando 5 della decisione).

Come poi si precisa nel considerando 4 della menzionata decisione: «L'Unione ha adottato norme comuni su cui incideranno gli elementi presi in esame in riferimento alla convenzione»; elementi che comprendono, in particolare, un insieme esaustivo di regole

⁷⁰ Già citata nella nota 20.

⁷¹ Regolamento (UE) 2024/1689, cit.

⁷² La proposta legislativa dell'AI ACT risale al 21 aprile 2021.

nel settore del mercato unico per i prodotti⁷³ e i servizi⁷⁴ per cui è possibile utilizzare i sistemi di IA, nonché il diritto derivato dell'Unione⁷⁵ che attua la Carta dei diritti fondamentali, considerato che in determinate circostanze lo sviluppo e l'utilizzo di taluni sistemi di IA potrebbero incidere negativamente su tali diritti.

Pertanto, prosegue la decisione nel considerando 6, «la conclusione della convenzione può incidere sulle norme comuni dell'Unione esistenti e su quelle

⁷³ Direttiva 85/374/CEE del Consiglio, del 25 luglio 1985, relativa al ravvicinamento delle disposizioni legislative, regolamentari ed amministrative degli Stati membri in materia di responsabilità per danno da prodotti difettosi, in GU L 210 del 7 agosto 1985, p. 29; direttiva 2001/95/CE del Parlamento europeo e del Consiglio, del 3 dicembre 2001, relativa alla sicurezza generale dei prodotti, in GU L 11 del 15 gennaio 2002, p. 4; direttiva 2006/42/CE del Parlamento europeo e del Consiglio, del 17 maggio 2006, relativa alle macchine e che modifica la direttiva 95/16/CE (rifusione), in GU L 157 del 9 giugno 2006, p. 24; direttiva 2009/48/CE del Parlamento europeo e del Consiglio, del 18 giugno 2009, sulla sicurezza dei giocattoli, in GU L 170 del 30 giugno 2009, p. 1; direttiva 2014/53/UE del Parlamento europeo e del Consiglio, del 16 aprile 2014, concernente l'armonizzazione delle legislazioni degli Stati membri relative alla messa a disposizione sul mercato di apparecchiature radio e che abroga la direttiva 1999/5/CE, in GU L 153 del 22 maggio 2014, p. 62; regolamento (UE) 2017/745 del Parlamento europeo e del Consiglio, del 5 aprile 2017, relativo ai dispositivi medici, che modifica la direttiva 2001/83/CE, il regolamento (CE) n. 178/2002 e il regolamento (CE) n. 1223/2009 e che abroga le direttive 90/385/CEE e 93/42/CEE del Consiglio, in GU L 117 del 5 maggio 2017, p. 1; regolamento (UE) 2019/2144 del Parlamento europeo e del Consiglio del 27 novembre 2019 relativo ai requisiti di omologazione dei veicoli a motore e dei loro rimorchi, nonché di sistemi, componenti ed entità tecniche destinati a tali veicoli, per quanto riguarda la loro sicurezza generale e la protezione degli occupanti dei veicoli e degli altri utenti vulnerabili della strada, che modifica il regolamento (UE) 2018/858 del Parlamento europeo e del Consiglio e abroga i regolamenti (CE) n. 78/2009, (CE) n. 79/2009 e (CE) n. 661/2009 del Parlamento europeo e del Consiglio e i regolamenti (CE) n. 631/2009, (UE) n. 406/2010, (UE) n. 672/2010, (UE) n. 1003/2010, (UE) n. 1005/2010, (UE) n. 1008/2010, (UE) n. 1009/2010, (UE) n. 19/2011, (UE) n. 109/2011, (UE) n. 458/2011, (UE) n. 65/2012, (UE) n. 130/2012, (UE) n. 347/2012, (UE) n. 351/2012, (UE) n. 1230/2012 e (UE) 2015/166 della Commissione, in GU L 325 del 16 dicembre 2019, p. 1.

⁷⁴ Direttiva 2000/31/CE del Parlamento europeo e del Consiglio dell'8 giugno 2000 relativa a taluni aspetti giuridici dei servizi della società dell'informazione, in particolare il commercio elettronico, nel mercato interno («direttiva sul commercio elettronico»), in GU L 178 del 17 luglio 2000, p. 1; direttiva 2006/123/CE del Parlamento europeo e del Consiglio del 12 dicembre 2006 relativa ai servizi nel mercato interno, in GU L 376 del 27 dicembre 2006, p. 36; direttiva 2008/48/CE del Parlamento europeo e del Consiglio, del 23 aprile 2008, relativa ai contratti di credito ai consumatori e che abroga la direttiva 87/102/CEE del Consiglio, in GU L 133 del 22 maggio 2008, p. 66.

⁷⁵ Direttiva 2000/43/CE del Consiglio, del 29 giugno 2000, che attua il principio della parità di trattamento fra le persone indipendentemente dalla razza e dall'origine etnica, in GU L 180 del 19 luglio 2000, p. 22; direttiva 2000/78/CE del Consiglio, del 27 novembre 2000, che stabilisce un quadro generale per la parità di trattamento in materia di occupazione e di condizioni di lavoro, in GU L 303 del 2 dicembre 2000, p. 16; Direttiva 2002/58/CE del Parlamento europeo e del Consiglio, del 12 luglio 2002, relativa al trattamento dei dati personali e alla tutela della vita privata nel settore delle comunicazioni elettroniche (direttiva relativa alla vita privata e alle comunicazioni elettroniche), in GU L 201 del 31 luglio 2002, p. 37; regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio, del 27 aprile 2016, relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati e che abroga la direttiva 95/46/CE (regolamento generale sulla protezione dei dati), in GU L 119 del 4.5.2016, p. 1; direttiva (UE) 2016/680 del Parlamento europeo e del Consiglio, del 27 aprile 2016, relativa alla protezione delle persone fisiche con riguardo al trattamento dei dati personali da parte delle autorità competenti a fini di prevenzione, indagine, accertamento e perseguimento di reati o esecuzione di sanzioni penali, nonché alla libera circolazione di tali dati e che abroga la decisione quadro 2008/977/GAI del Consiglio, in GU L 119 del 4 maggio 2016, p. 89).

prevedibilmente elaborate in futuro o modificarne la portata ai sensi dell'articolo 3, paragrafo 2, del trattato sul funzionamento dell'Unione europea (TFUE)»⁷⁶.

Alla stregua di tali considerazioni l'Unione ha, pertanto, contribuito alla conclusione della Convenzione per quanto concerne le materie rientranti nella sua competenza esclusiva, al precipuo fine di tutelare l'integrità del diritto dell'Unione e di garantire che le norme del diritto internazionale e unionale rimassero coerenti⁷⁷. Ciò nella convinzione che la Convenzione potesse stabilire norme internazionali di alto livello in materia di regolamentazione dell'IA con un impatto sui diritti umani, sul funzionamento della democrazia e sul rispetto dello Stato di diritto, specie alla luce dei lavori già svolti dal Consiglio d'Europa in tale settore⁷⁸.

Sul punto, anche il Garante europeo della protezione dei dati, nell'esprimere la sua opinione al riguardo⁷⁹, aveva dichiarato che la Convenzione rappresentasse un'importante opportunità per integrare l'allora proposta di legge sull'IA, rafforzando in tal modo la protezione dei diritti fondamentali di tutte le persone interessate dai sistemi di IA.

Tanto premesso, la Convenzione *de qua* contiene, dunque, disposizioni specifiche relative all'Unione europea, che sono senza dubbio il risultato del coinvolgimento della Commissione UE all'interno dei negoziati.

Anzitutto, nel Capitolo VIII relativo alle clausole finali, l'art. 27 rubricato *Effetti della Convenzione*, al par. 2 precisa gli Stati-parte che sono membri dell'Unione europea applicheranno, nelle loro relazioni reciproche, le norme dell'UE che disciplinano le questioni rientranti nell'ambito di applicazione della Convenzione, ma con un limite: senza pregiudicare l'oggetto e lo scopo della Convenzione, nonché la sua piena applicazione con altre Parti.

Al riguardo, come si chiarisce nell'*Explanatory Report*, il paragrafo 2 dell'art. 27 riconosce «the increased integration of the European Union, particularly as regards regulation of artificial intelligence systems» e consente ai Paesi UE di applicare tra di essi e, tra questi ultimi e le istituzioni, gli organi, gli uffici e le agenzie dell'Unione europea,

⁷⁶ J. ZILLER, *The Council of Europe Framework Convention on Artificial Intelligence vs. the EU Regulation: two quite different legal instruments*, cit., p. 206, ove l'A. ritiene che «It is quite striking that this recital refers to Article 3(2) on the values of the Union, according to which 'The Union shall offer its citizens an area of freedom, security and justice without internal frontiers, in which the free movement of persons is ensured in conjunction with appropriate measures with respect to external border controls, asylum, immigration and the prevention and combating of crime'». A tal proposito precisa, infatti, che «the Commission proposal of 21 April 2021, as well as the text approved by the European Parliament in March 2024 refer only to the legal bases relating to the internal market and not to those relating to controls at the internal and external borders of the Union (Article 77(2) TFEU).

⁷⁷ Considerando n. 7 della decisione.

⁷⁸ Considerando n. 8 della Convenzione.

⁷⁹ Parere 20/2022 del Garante europeo della protezione dei dati sulla raccomandazione di decisione del Consiglio che autorizza l'avvio di negoziati a nome dell'Unione europea per una convenzione del Consiglio d'Europa sull'intelligenza artificiale, i diritti umani, la democrazia e lo Stato di diritto. Tale parere del GEPD è emesso in risposta a una consultazione della Commissione del 18 agosto 2022 ai sensi dell'art. 42, par. 1, del Regolamento (UE) 2018/1725 del Parlamento europeo e del Consiglio, del 23 ottobre 2018, sulla tutela delle persone fisiche in relazione al trattamento dei dati personali da parte delle istituzioni, degli organi e degli organismi dell'Unione e sulla libera circolazione di tali dati, e che abroga il regolamento (CE) n. 45/2001 e la decisione n. 1247/2002/CE, in GU L 295 dell'21 novembre 2018, p. 39.

il DUE che disciplina le medesime questioni trattate nella Convenzione quadro, ove per “diritto” si intende in senso ampio «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»⁸⁰.

Lo stesso vale per le altre Parti che applicano le norme dell'Unione europea nella misura in cui sono da queste vincolate, in considerazione della loro partecipazione al mercato interno dell'UE o dell'essere soggette al trattamento del mercato interno medesimo.

In ogni caso, il menzionato art. 27 garantisce che non sia pregiudicata la piena applicazione della Convenzione quadro tra l'Unione europea o le Parti che sono membri dell'Unione europea e le altre Parti, nonché tra Parti che non sono membri dell'Unione europea nella misura in cui sono anch'esse vincolate dalle stesse norme e altre Parti della Convenzione quadro.

Ne discende, dunque, che, una volta ratificata la Convenzione dall'Unione, gli Stati membri potranno, ad esempio, denunciare le violazioni del DUE che costituiscano anche una violazione della Convenzione, nella misura in cui sono vincolate da tali norme.

E ancora, l'Unione europea viene menzionata, altresì, all'art. 28 rubricato “modifiche”, ove si specifica che gli emendamenti alla Convenzione possono essere proposti da qualsiasi Parte, così come dal Comitato dei Ministri dell'Unione europea, oltre che dalla Conferenza delle Parti; al riguardo, come puntualizza il paragrafo 149 dell'*Explanatory Report*, per “emendamenti” si intendono principalmente modifiche minori di carattere procedurale e tecnico, atteso che «major changes to the Framework Convention could be made in the form of amending protocols».

Infine, ai sensi dell'art. 30, par. 1, la Convenzione è aperta alla firma, altresì, dell'Unione europea, la quale è stata apposta in data 5 settembre 2024, suggellandone il ruolo di Parte contraente della Convenzione medesima.

A tal proposito, considerata l'attiva partecipazione dell'Unione europea nei negoziati, con l'inserimento di disposizioni *ad hoc* nel testo convenzionale, che depongono per un testo “concordato” tra UE e Consiglio d'Europa, potrebbe ipotizzarsi una ratifica a breve da parte dell'Unione secondo l'*iter* di cui all'art. 218 TFUE⁸¹.

⁸⁰ *Explanatory Report*, cit., par. 147.

⁸¹ Sulla base di una raccomandazione della Commissione o dell'Alto rappresentante, il Consiglio autorizza l'avvio dei negoziati, nomina il negoziatore dell'Unione, adotta le direttive di negoziato, autorizza la firma degli accordi e li conclude. Il Consiglio delibera a maggioranza qualificata per tutta la durata della procedura, oppure all'unanimità quando l'accordo riguarda un settore per il quale è richiesta l'unanimità per l'adozione di un atto dell'Unione. In pratica, tende a cercare l'unanimità anche quando non è richiesta dalla legge. L'approvazione del Parlamento europeo è richiesta in particolare per l'adesione dell'Unione alla convenzione europea per la salvaguardia dei diritti dell'uomo e delle libertà fondamentali e per “accordi che riguardano settori ai quali si applica la procedura legislativa ordinaria oppure la procedura legislativa speciale qualora sia necessaria l'approvazione del Parlamento europeo”. Cfr. G. ZACCARONI, *Intelligenza artificiale e principio democratico: riflessioni a margine dell'emersione di un quadro normativo europeo*, cit., p. 36. L'A. evidenzia che, in questo caso, non sembrano trovare applicazione le disposizioni dell'art. 218 TFUE che rafforzano la procedura di conclusione degli accordi internazionali nel caso in cui si tratti di aderire alla CEDU, come la previa approvazione del Parlamento europeo, atteso che quest'ultima è menzionata espressamente, cosa che non avviene invece riguardo la possibilità di entrare a fare parte di

Del resto, l'UE è già parte di convenzioni, come la Convenzione delle Nazioni Unite sui diritti delle persone con disabilità (CRPD)⁸², ratificata dall'UE il 23 dicembre 2010⁸³ o la Convenzione del Consiglio d'Europa sulla prevenzione e la lotta contro la violenza nei confronti delle donne e la violenza domestica (Convenzione di Istanbul), ratificata il 1° giugno 2023 con due decisioni del Consiglio⁸⁴.

In tale ambito, rilevante indubbiamente sarà l'interazione tra la Corte di giustizia e la Corte europea dei diritti dell'uomo, nonché l'influenza della giurisprudenza interpretativo-creativa di quest'ultima.

Al riguardo valga menzionare la sentenza *Ships Waste Oil Collector B.V. e altri c. Paesi Bassi* del 1° aprile 2025⁸⁵, ove la Corte europea nel ragionare sulla nozione di vittima e, in particolare, sull'ammissibilità *ratione personae* di ricorsi individuali presentati da società commerciali ex art. 34 CEDU, si sofferma su un'interessante discussione, seppur "theoretical and philosophical"⁸⁶ relativa all'intelligenza artificiale.

Come noto, l'art. 34 consente la legittimazione a presentare ricorso individuale a "una persona fisica", "un'organizzazione non governativa" o "un gruppo di privati". Ebbene, osserva la Corte, nonostante la classificazione delle società in una qualsiasi delle tre "windows" or "doorways" del citato articolo 34 sia problematica, la soluzione meno difficoltosa è quella delle "organizzazioni non governative", da intendere, dunque, come comprensive, altresì, delle società. Esclusa ovviamente la riconducibilità delle società nei "gruppi di privati" – rileva la Corte – sarebbe più difficile per le società passare attraverso la "finestra" delle "persone", specie se si consideri l'espressione "*personne physique*" nel testo francese. E sulla base di tali considerazioni, al par. 47 e ss., la Corte rileva che, non solo il termine francese impedisce alle società di avere lo *status* di vittima attraverso la "finestra" delle "persone", ma potrebbe anche impedire alla Corte, nella sua futura giurisprudenza, di attribuire lo *status* di vittima a diverse forme di intelligenza artificiale, specie se quest'ultime, che possiedono alcune delle funzioni proprie delle persone fisiche, come i robot, chiedano protezione alla Corte⁸⁷.

Così ragionando sul punto, al paragrafo 49 la Corte asserisce che, se un giorno gli Stati membri decidessero di garantire una qualche forma di protezione all'intelligenza artificiale, sarebbe più opportuno farlo attraverso un nuovo Protocollo, il quale offrirebbe, altresì, l'opportunità di stabilire con fermezza e chiarezza una base giuridica per le rivendicazioni di entità diverse dalle persone fisiche, comprese le società. Tuttavia, la

altre Convenzioni previste dal Consiglio d'Europa. Qualora dunque l'Unione decida di stipulare la Convenzione quadro sull'intelligenza artificiale, la procedura per la sua conclusione potrebbe paradossalmente terminare prima di quella per l'adesione alla CEDU.

⁸² Convention on the Rights of Persons with Disabilities (CRPD).

⁸³ Ratifica avvenuta con la decisione del Consiglio dell'UE 2010/48/CE. L'UE è l'unica organizzazione regionale ad aver ratificato la CRPD.

⁸⁴ Si ricorda che l'1 giugno 2023, l'Unione europea, con due decisioni del Consiglio (la n. 2023/1075 e la n. 2023/1076) ha aderito alla Convenzione del Consiglio d'Europa sulla prevenzione e la lotta alla violenza contro le donne e la violenza domestica (Convenzione di Istanbul).

⁸⁵ Corte europea dei diritti dell'uomo, sentenza dell'1 aprile 2025, ricorsi nn. 2799/16 e altri 3, *Ships Waste Oil Collector B.V. e altri c. Paesi Bassi*.

⁸⁶ Corte europea dei diritti dell'uomo, sentenza *Ships Waste Oil Collector B.V. e altri*, cit., par. 49.

⁸⁷ Corte europea dei diritti dell'uomo, sentenza *Ships Waste Oil Collector B.V. e altri*, cit., par. 48.

Corte si chiede pure come simili, eventuali sviluppi futuri potrebbero mai conciliarsi con l'obiettivo primario della Convenzione, quale quello di fornire un'efficace tutela dei diritti umani e del principio della dignità della persona. E qui conclude affermando che, se in futuro l'intelligenza artificiale dovesse godere di “similar or adjacent to human rights” tutelati dalla Convenzione, come garantito da un nuovo Protocollo, è fondamentale che ciò vada di pari passo con la capacità e l'obbligo di tale intelligenza artificiale di rispettare i diritti umani, lo Stato di diritto e il principio di democrazia⁸⁸.

L'*Explanatory Report*, tra l'altro, con riferimento all'obbligo per le Parti della Convenzione di assicurare che tutte le attività che fanno parte del ciclo di vita dell'intelligenza artificiale siano conformi agli obblighi di tutela dei diritti fondamentali, sia a livello internazionale che a livello interno, non manca di elencare puntualmente le fonti pertinenti di diritto dell'Unione europea, tra cui la Carta dei diritti fondamentali.

Inoltre, degno di menzione è il paragrafo 83 dell'*Explanatory Report* ove, con riferimento all'art. 11 in tema di *Privacy* e protezione dei dati personali, si sottolinea come, a livello dell'UE, il Regolamento generale sulla protezione dei dati 2016/679, “GDPR” sia «a comprehensive data protection law» che si applica «to natural or legal persons that process personal data belonging to natural persons in the European Union regardless of whether the processing takes place in the European Union or not», riconoscendone la vasta portata applicativa.

In definitiva, da questa sintetica lettura di disposizioni della Convenzione e delle relative spiegazioni, emerge un palese “ravvicinamento” tra Consiglio d'Europa e Unione europea che lascia trasparire possibili progressi anche rispetto all'*iter* di adesione dell'Unione europea alla CEDU⁸⁹, specie considerato il substrato di valori, principi e diritti che accomuna sia l’“intelligenza naturale/umana” che “artificiale”.

⁸⁸ L'intelligenza artificiale viene menzionata, altresì, nella Partly Dissenting Opinion of Judge Pavli, relativa alla sentenza della Corte europea, del 4 giugno 2019, ricorso n. 39757/15, *Sigurður Einarsson e altri c. Irlanda*, punto 15: «Emerging practice in the Council of Europe area is in line with this general approach. Thus, courts in at least two jurisdictions (the United Kingdom and Ireland) have approved in recent years the use of technology-assisted review, employing a form of artificial intelligence known as predictive coding, for the purposes of electronic disclosure in high-stakes civil litigation. The rationale would apply with equal force in criminal cases of comparable complexity. Again, the underlying premise for the use of such advanced technology is, of course, that both sides are granted the fullest possible access to begin with. And, secondly, that criminal-law frameworks and investigative practices are organised in such a way as to facilitate adequate access for the defence at the appropriate (that is, early) stage of proceedings»; punto 21: «The same would apply to the second justification put forward by the majority, namely that the defence failed to specify what kind of additional evidence they were seeking, relying for example on “overviews of the seized items and their rough content” (see paragraph 92 in fine). With respect, this argument severely underestimates, in my view, the complexities of analysing large and interconnected amounts of investigative data, whether one is equipped with “merely” human intelligence or aided by artificial intelligence». Enfasi aggiunta.

⁸⁹ Dopo la fase di stallo nell'*iter* di adesione dell'Unione europea alla CEDU, determinata dal parere 2/13 emesso il 18 dicembre 2014, con cui la Corte di giustizia si è pronunciata sulla incompatibilità del progetto di accordo di adesione rispetto al diritto dell'Unione europea, a distanza di anni, il gruppo ad hoc (“46+1”) ha tenuto la sua 18ª riunione in presenza dal 14 al 17 marzo 2023, al fine di negoziare un testo che provasse a superare le censure mosse al progetto di accordo nel citato parere 2/13. A tal riguardo il gruppo ha concluso di aver risolto tutte le questioni che avrebbe dovuto affrontare, prendendo atto, invece, dell'intenzione dell'UE di risolvere internamente la questione in sospeso relativa agli atti dell'UE nell'ambito della politica estera e di sicurezza comune, a cui il gruppo “46+1” non è riuscito a trovare una

9. Considerazioni conclusive sul *quid pluris* della Convenzione quadro

Come emerso, il crescente utilizzo dei sistemi di intelligenza artificiale è foriero di nuove sfide per i diritti umani, la democrazia e lo Stato di diritto.

Ad oggi, in ambito internazionale, europeo e regionale, il Consiglio d'Europa è l'unico ad aver adottato un atto pattizio in materia, fornendo criteri guida per orientare programmatori e utilizzatori dei sistemi di intelligenza artificiale in ogni fase del loro *c.d. lifecycle*.

A tal uopo la Convenzione quadro si pregia di un doppio valore aggiunto. Anzitutto, come evidenziato in premessa, è un "trattato aperto", con una portata potenzialmente globale/universale: tutti i paesi, anche non europei, a prescindere da se siano Stati-parte del Consiglio d'Europa e indipendentemente dal fatto che abbiano partecipato ai lavori di stesura e negoziazione della Convenzione del CAI, potranno aderirvi e vincolarsi al rispetto degli standard ivi stabiliti. Siffatta Convenzione trascende, dunque, i confini "europei", coinvolgendo e attraendo non solo i Paesi d'Europa, bensì il maggior numero possibile di Stati di tutte le regioni del mondo.

Al riguardo, vero è che le disposizioni convenzionali prevedono obblighi in parte sovrapponibili a quelli previsti dal Regolamento UE 2024/1689 e che quest'ultimo sia direttamente applicabile, ma, nel caso della Convenzione, proprio in ragione della sua destinatarietà, esse assumono un campo applicativo ben più vasto, vincolando non soltanto gli Stati membri dell'Unione ma potenzialmente tutti i Paesi.

In secondo luogo, la Convenzione, al pari del Regolamento UE sull'IA, si fonda su una metodologia di rischio (approccio *risk-based*), la quale, in special modo per quanto concerne la Convenzione, guida e assiste nell'identificazione di contesti e applicazioni rispetto alle quali l'implementazione di sistemi di intelligenza artificiale potrebbe presentare rischi non solo per il godimento dei diritti umani, bensì anche per il funzionamento della democrazia e l'osservanza dello Stato di diritto. In particolare, essa impone una gestione preventiva dei potenziali impatti negativi, sollecitando gli Stati a vietare l'uso di tecnologie incompatibili con i diritti fondamentali.

A tal proposito valga menzionare la metodologia denominata HUDERIA⁹⁰, quale acronimo di *Human Rights, Democracy and Rule of Law Impact Assessment*. Sebbene

soluzione. Su tale base, il citato gruppo ha adottato la relazione contenente il nuovo pacchetto di strumenti di adesione rivisti, comprensivo delle bozze definitive dell'accordo e del rapporto esplicativo che lo accompagna (<https://rm.coe.int/steering-committee-for-human-rights-cddh-interim-report-to-the-committ/1680aace4e>). Cfr. N. BERGAMASCHI, *Prime considerazioni sul nuovo tentativo di adesione dell'Unione alla CEDU e sui suoi principali ostacoli*, in *Quaderni AISDUE*, 2024, www.aisdue.eu/nicola-bergamaschi-prime-considerazioni-sul-nuovo-tentativo-di-adesione-dellunione-alla-cedu-e-sui-suoi-principali-ostacoli/.

⁹⁰ Committee on Artificial Intelligence (CAI), *Methodology for the Risk and Impact Assessment of Artificial Intelligence Systems from the Point of View of Human Rights, Democracy and the Rule of Law (Huderia Methodology)*, Strasbourg, 28 November 2024, <https://rm.coe.int/cai-2024-16rev2-methodology-for-the-risk-and-impact-assessment-of-arti/1680b2a09f>. La Metodologia HUDERIA è stata adottata dal Comitato sull'intelligenza artificiale (CAI) del Consiglio d'Europa durante la sua 12a sessione plenaria, tenutasi a Strasburgo dal 26 al 28 novembre 2024. Può essere utilizzata da attori sia pubblici che privati per identificare e affrontare i rischi e l'impatto sui diritti umani, sulla democrazia e sullo Stato di diritto

essa non sia di natura giuridicamente vincolante ed esistano già altre metodologie di rischio e impatto sui diritti umani, HUDERIA presenta un valore aggiunto, in quanto non soltanto combina per la sua progettazione gli aspetti tecnici dell'IA insieme al contesto socio-tecnico del suo sviluppo e applicazione, ma incorpora anche l'intero *acquis* del Consiglio d'Europa non solo in termini di diritti umani, ma altresì di democrazia e Stato di diritto.

Ne emerge, dunque, l'approccio fortemente umano-centrico⁹¹, volto a prevenire e ridurre i potenziali rischi che possono emergere dall'uso e dall'evolversi dell'IA, affinché quest'ultima sia al servizio dell'uomo e non lo sovrasti, a differenza della regolamentazione UE, più protesa verso la tutela del diritto alla riservatezza e alla protezione dei dati.

Le componenti chiave di tale approccio possono ravvisarsi in una serie di meritori aspetti valorizzati dalla Convenzione, e sopra passati in rassegna, come la partecipazione attiva di utenti e parti interessate, coinvolti direttamente nel processo di design e valutazione dei sistemi IA, il che include – si ribadisce – la consultazione di rappresentanti della società civile, esperti di etica, giuristi, e simili, per assicurare che le applicazioni IA riflettano un ampio spettro di interessi umani; la protezione della dignità umana, che è uno dei pilastri dell'approccio umano-centrico, il che implica che i sistemi IA siano progettati in modo da non deumanizzare gli individui, né ridurli a semplici dati⁹²; la minimizzazione dei rischi, prevenendo situazioni in cui l'intelligenza artificiale possa causare danni involontari e ciò include lo sviluppo di meccanismi di allerta precoce e sistemi di revisione per identificare e mitigare possibili impatti negativi prima che si manifestino; l'equità e la non discriminazione, tali per cui i sistemi IA debbano essere esenti da pregiudizi, non perpetuino disparità esistenti, né introducano nuove forme di disuguaglianza, richiedendo una rigorosa verifica e validazione delle tecnologie; la trasparenza, che mette in risalto la necessità che i sistemi AI siano chiari e accessibili agli utenti, consentendo loro di comprendere le decisioni prese ed eventualmente contestarle.

Alla stregua di tale approccio, la Convenzione *de qua*, di carattere trasversale, si pone, in definitiva, in armonia con i sistemi e meccanismi di protezione dei diritti umani esistenti, applicabili da ciascuna Parte in ottemperanza ai vigenti obblighi sia di diritto

nell'intero ciclo di vita dei sistemi di IA. Essa prevede la creazione di un piano di mitigazione dei rischi per ridurre al minimo o eliminare i rischi individuati, proteggendo le persone da potenziali danni. Se si constata, ad esempio, che un sistema di IA utilizzato in ambito di assunzioni mostra pregiudizi nei confronti di determinati gruppi demografici, il piano di mitigazione potrebbe prevedere la correzione dell'algoritmo o l'attuazione di una revisione umana. La metodologia richiede rivalutazioni periodiche per assicurare che il sistema di IA continui a operare in modo sicuro ed etico di pari passo con l'evoluzione del contesto e della tecnologia. Questo approccio garantisce la protezione delle persone da rischi emergenti in tutto il ciclo di vita del sistema di IA.

⁹¹ Si segnala, *inter alia*, la Dichiarazione di Bletchley, sottoscritta in occasione dell'AI Safety Summit, da 28 Paesi. Si tratta di un accordo in cui gli stati firmatari si sono impegnati in una cooperazione internazionale finalizzata al raggiungimento di un'IA umano-centrica.

⁹² Al riguardo si segnala: UNESCO Recommendation on the Ethics of Artificial Intelligence, adopted on 23 November 2021. La protezione dei diritti umani e della dignità è la pietra angolare della Raccomandazione, basata sul progresso di principi fondamentali come la trasparenza e l'equità, ricordando sempre l'importanza della supervisione umana dei sistemi di IA.

internazionale che nazionale. Essa, infatti, non intende creare nuovi diritti o obblighi in materia di diritti umani o minare la portata e il contenuto degli strumenti di protezione applicabili, ma piuttosto facilitarne l'attuazione nel contesto delle nuove sfide globali sollevate dall'intelligenza artificiale. A tal uopo essa pone al centro la persona, segnando una significativa svolta nella direzione di "umanizzare" l'IA e modellare l'evoluzione tecnologica in una prospettiva "etica" e responsabile, a garanzia dei diritti fondamentali, della democrazia e della *rule of law*.

ABSTRACT: Il presente contributo intende fornire un'analisi critica e puntuale della Convenzione quadro sull'intelligenza artificiale e i diritti umani, la democrazia e lo Stato di diritto, adottata il 17 maggio 2024 dal Comitato dei Ministri del Consiglio d'Europa, evidenziando talora profili di affinità o divergenza rispetto al coevo Regolamento (UE) 2024/1689 del Parlamento europeo e del Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale, meglio conosciuto come *AI Act* (*Artificial Intelligence Act*), per poi concludere con qualche riflessione sul *quid pluris* del testo convenzionale.

KEYWORDS: Convenzione quadro Consiglio d'Europa – Intelligenza artificiale – Diritti umani – Democrazia – Stato di diritto.

COUNCIL OF EUROPE AND ARTIFICIAL INTELLIGENCE: A FIRST ATTEMPT AT REGULATION TO PROTECT HUMAN RIGHTS, DEMOCRACY AND THE RULE OF LAW

ABSTRACT: This contribution aims to provide a critical and timely analysis of the Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, adopted on 17 May 2024, by the Committee of Ministers of the Council of Europe, highlighting at times profiles of affinity or divergence with respect to the contemporary Regulation (EU) 2024/1689 of the European Parliament and of the Council laying down harmonized rules on artificial intelligence, better known as the AI Act (Artificial Intelligence Act), and finally concluding with some reflections on the *quid pluris* of the conventional text.

KEYWORDS: Council of Europe Framework Convention – Artificial Intelligence – Human Rights – Democracy – Rule of Law.