SIOI
UNA Italy
OSSERVATORIO
sulle attività delle organizzazioni internazionali e sovranazionali, universali e regionali, sui temi di interesse della politica estera italiana

La protezione dei diritti fondamentali nel regolamento dell'Unione europea sull'intelligenza artificiale

Mariaida Cristarella Oristano
Dottoranda di ricerca in Diritto pubblico, internazionale ed europeo, Università degli Studi di Milano La Statale

In una fase storica in cui l'esigenza di migliorare il benessere sociale attraverso il progresso tecnologico si contrappone ai numerosi risvolti negativi derivanti da un uso improprio dello stesso, è fondamentale promuovere l'innovazione garantendo un elevato livello di tutela dei diritti umani[1]. Tale necessità sembra ancora più urgente se si considera il settore della c.d. "intelligenza artificiale" (IA) che, consentendo la programmazione e la progettazione di sistemi *hardware* e *software* capaci di conferire alle macchine qualità umane[2], è divenuta una delle tecnologie più strategiche e, al contempo, controverse del secolo[3]. Infatti, se i benefici dell'adozione dell'IA per i cittadini, le imprese e il commercio sono ampiamente riconosciuti[4], alcune sue applicazioni sollevano numerosi interrogativi etici e giuridici in termini di responsabilità, concorrenza, sicurezza, trasparenza e finanche tutela dei diritti fondamentali[5].

In tal contesto, l'Unione europea (UE) riveste una posizione particolarmente importante, promuovendo una progettazione ed un'utilizzazione dell'IA tale da migliorare la società senza che siano sacrificati i diritti dei cittadini[6]. Infatti, l'UE – dopo aver istituito un gruppo indipendente di esperti ad alto livello[7] e rilevato l'inadeguatezza

[1] Con riferimento agli effetti della tecnologia sui diritti dell'uomo, v. L. PANELLA, *Nuove tecnologie e diritti umani: profili di diritto internazionale e di diritto interno*, Napoli, 2018; O. POLLICINO., *Tutela dei diritti fondamentali nell'era digitale e contesto valoriale: una indagine transatlantica*, in *Rivista MediaLaws*, 2018, pp. 39-52; ID., *Judicial Protection of Fundamental Rights on the Internet. A Road Towards Digital Constitutionalism?*, London, 2022; M. LAND, J. ARONSON (eds.), *Normative Approaches to Technology and Human Rights*, in ID., *New Technologies for Human Rights Law and Practice*, Cambridge, 2019, pp. 21-124; M. OLIVETTI, *Diritti fondamentali e nuove tecnologie*, in *Revista Estudos Institucionais*, 2020, pp. 395-430.

[2] Tra queste, l'apprendimento, il ragionamento, la creatività e la pianificazione. In altri termini, «l'intelligenza artificiale permette ai sistemi di capire il proprio ambiente, mettersi in relazione con quello che percepisce e risolvere problemi, e agire verso un obiettivo specifico» (*Che cos'è l'intelligenza artificiale e come viene usata?* in https://www.europarl.europa.eu/news/it/headlines/priorities/intelligenza-artificiale-nell-ue/20200827STO85804/che-cos-e-l-intelligenza-artificiale-e-come-viene-usata).

[3] Comunicazione della Commissione al Parlamento europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle regioni del 25 aprile 2018, *L'intelligenza artificiale per l'Europa*, p. 2.

[4] Per esempio, nel suo Quadro di valutazione della trasformazione digitale del 2018 (disponibile in https://ati.ec.europa.eu/), la Commissione europea ha evidenziato che le imprese dei settori edile ed agroalimentare che hanno fatto ricorso all'IA confermano effetti positivi in termini di migliori servizi e prodotti, acquisizione di nuovi clienti ed ingresso in nuovi mercati.

[5] In argomento, la bibliografia è vasta: cfr., in particolare, A. AMIDEI, Intelli*genza artificiale e diritti umani: prospettive e limiti delle tecnologie di potenziamento*, in *I diritti dell'uomo*, 2020, pp. 361-384; G. SARTOR, *Artificial Intelligence and Human Rights: Between Law and Ethics*, in *Maastricht Journal of European and Comparative Law*, 2020, pp. 705-719; C. GRIECO, *Intelligenza artificiale e diritti umani nel diritto internazionale e dell'Unione europea. Alla ricerca di un delicato equilibrio*, in *Rivista OIDU*, 2022, pp. 782-810.

[6] Cfr. F. FERRI (a cura di), *L'Unione europea e la nuova disciplina sull'intelligenza artificiale: questioni e prospettive*, Napoli, 2024.

[7] Gruppo di esperti ad alto livello sull'intelligenza artificiale che, già nell'aprile 2019, ha pubblicato gli orientamenti etici per lo sviluppo e l'utilizzo dell'intelligenza artificiale (*Orientamenti etici per un'IA affidabile*, 8 aprile 2019).

della normativa vigente di fronte ai problemi legati all'uso dell'IA[8] – ha di recente adottato il primo atto legislativo al mondo che disciplina la materia identificandone i rischi e definendone le misure di protezione, con particolare attenzione alla tutela dell'individuo e dei suoi diritti fondamentali. Si sta facendo riferimento al regolamento sull'intelligenza artificiale, anche noto come "*Artificial Intelligence Act*" (d'ora in poi *AI Act*)[9], uno strumento orizzontale che persegue il duplice obiettivo di creare un mercato unico per l'IA che fornisca regole chiare e uniformi per gli operatori del settore[10]; e promuovere uno sviluppo tecnologico orientato alla tutela dell'individuo, attraverso l'introduzione di obblighi per gli sviluppatori di IA e standard tecnici che garantiscano l'accesso al mercato europeo a quei soli sistemi ritenuti affidabili e conformi ai valori dell'UE e ai diritti umani[11].

Tali obiettivi si riflettono nella doppia base giuridica del regolamento: l'art. 114 TFUE, che permette l'adozione di misure per armonizzare le normative nazionali riguardanti il funzionamento del mercato interno; e l'art. 16 TFUE, che conferisce all'Unione il potere di emanare normative sulla tutela dei dati personali. Si tratta di un approccio adottato, peraltro, anche nell'ambito del Consiglio d'Europa, in cui il Comitato dei ministri, nella 133ª sessione ministeriale del 17 maggio 2024, ha approvato all'unanimità la Convenzione sull'intelligenza artificiale[12], finalizzata a garantire che l'impiego dei sistemi di IA avvenga nel rispetto dei principi fondamentali in materia di diritti umani, democrazia e Stato di diritto. Così, sebbene il regolamento in studio sia entrato in vigore il 2 agosto 2024 e le sue disposizioni saranno pienamente efficaci a partire dal 2 agosto 2026[13], alcune norme (e in particolare i capi I e II relativi alle disposizioni generali e alle pratiche di IA vietate), sono divenute applicabili già a decorrere dal 2 febbraio 2025, proprio in ragione dell'elevato pericolo che certi utilizzi di questo strumento rappresentano per i diritti delle persone. Invero – alla luce dei suddetti obiettivi e come indicato nella relativa proposta della Commissione – l'atto distingue sì vari livelli di rischio associato agli usi specifici dell'IA, definendo per ognuno di essi regole differenti[14], ma concentra la maggior parte delle sue disposizioni sulle applicazioni di tale settore che costituiscono una minaccia inaccettabile o alta per i diritti umani. Quanto alle prime, il regolamento proibisce, ad esempio, l'uso di tecniche subliminali

[8] Libro bianco sull'intelligenza artificiale – Un approccio europeo all'eccellenza e alla fiducia, 19 febbraio 2020, COM(2020) 65 final.

[9] Regolamento (UE) 2024/1689 del Parlamento europeo e del Consiglio, del 13 giugno 2024. In argomento, v. M. CARTA, *Il Regolamento UE sull'Intelligenza Artificiale: alcune questioni aperte*, in *Rivista Eurojus*, 2024, p. 188 ss.; M. INGLESE, *Il regolamento sull'intelligenza artificiale come atto per il completamento e il buon funzionamento del mercato interno?*, in *Rivista Quaderni AISDUE*, 2024, p. 71 ss.; A. CORRERA, *Il ruolo dell'Intelligenza artificiale nel paradigma europeo dell'E-justice. Prime riflessioni alla luce dell'AI Act*, ivi, p. 209 ss.

[10] *AI Act*, considerando n. 5.

[11] *AI Act*, considerando n. 1.

[12] Convenzione aperta alla firma il 5 settembre 2024.

[13] *AI Act*, art. 113.

[14] In particolare, si fa riferimento ai c.d. rischi minimi o nulli, limitati, elevati e inaccettabili. I primi riguardano tecnologie di IA, come i filtri *antispam*, cui non si applicano le norme dell'atto; i secondi interessano applicazioni che devono rispettare obblighi di trasparenza, come le *chatbot*; infine, gli ultimi si riferiscono a sistemi che, minacciando i diritti o la sicurezza, sono soggetti a normative stringenti (rischi elevati) o addirittura vietati (rischi inaccettabili). Inoltre, l'atto introduce anche la nozione di "rischio sistemico", definito all'art. 3(65) come «un rischio specifico per le capacità di impatto elevato dei modelli di IA per finalità generali, avente un impatto significativo sul mercato dell'Unione a causa della sua portata o di effetti negativi effettivi o ragionevolmente prevedibili sulla salute pubblica, la sicurezza, i diritti fondamentali o la società nel suo complesso, che può propagarsi su larga scala lungo l'intera catena del valore».

finalizzate a manipolare le persone senza il loro consenso o consapevolezza[15]; l'impiego di sistemi di IA che sfruttino le vulnerabilità di gruppi particolarmente esposti, come minori o persone con disabilità, alterandone significativamente il comportamento in modo tale da causare loro (o a terzi) un danno psico-fisico; o, ancora, l'uso di questa tecnologia per la c.d. "identificazione biometrica remota in tempo reale"[16] in spazi accessibili al pubblico a fini di attività di contrasto[17], fatte salve alcune eccezioni[18].

Per quanto riguarda, invece, i modelli di IA ad alto rischio nel mercato dell'UE, l'atto impone che tali sistemi – che devono assicurare un adeguato livello di accuratezza, robustezza e cibersicurezza (art. 15) – siano sottoposti a un processo di marcatura e valutazione volto a dimostrare la loro conformità a determinati requisiti obbligatori che ne garantiscano l'affidabilità. In particolare, sono stabiliti doveri in materia di *governance* dei dati (art. 10), documentazione (art. 11), conservazione delle registrazioni (art. 12), trasparenza e comunicazione delle informazioni agli utenti (art. 13). Inoltre, l'art. 14 dell'*AI Act* prevede che i sistemi di IA ad alto rischio siano progettati e sviluppati anche attraverso adeguati strumenti di interfaccia uomo-macchina, in modo da garantire una supervisione efficace da parte di operatori umani durante il loro utilizzo (art. 14). La responsabilità di verificare la conformità delle tecnologie di IA ai suddetti *standard* incombe primariamente sui fornitori, che devono svolgere verifiche interne. Tuttavia, su richiesta della Commissione, tali soggetti possono dover ricorrere a forme esterne di convalida, come i meccanismi di certificazione o la conformità a standard tecnici armonizzati elaborati dagli organismi di normazione per garantire il pieno rispetto dei requisiti legali.

Come è possibile osservare, in virtù dell'obbligo per il legislatore europeo di operare nel pieno rispetto dei diritti fondamentali, l'*AI Act* riserva ampio spazio alla loro tutela, in linea con le precedenti normative orizzontali in materia di sicurezza dei prodotti che includono, tra i propri obiettivi, anche la salvaguardia di tali diritti[19]. Conformemente a queste normative, infatti, anche il regolamento in studio prevede la c.d. *Fundamental Rights Impact Assessment* (FRIA), che impone agli sviluppatori dei sistemi di IA classificati come ad alto rischio di condurre, prima dell'utilizzo dei modelli, una

[15] *AI Act*, art. 5, lett. a).

[16] Ai sensi dell'art. 3(42), un «sistema di identificazione biometrica remota in tempo reale è un sistema di IA utilizzato per identificare persone fisiche senza il loro coinvolgimento attivo, generalmente a distanza, confrontando i dati biometrici di una persona con quelli contenuti in una banca dati di riferimento. Tale sistema deve operare senza ritardi significativi, includendo sia le identificazioni immediate che quelle con brevi ritardi, limitati a evitare l'elusione del sistema».

[17] Sono "attività di contrasto" quelle «attività svolte dalle autorità di contrasto o per loro conto a fini di prevenzione, indagine, accertamento o perseguimento di reati o esecuzione di sanzioni penali, incluse la salvaguardia contro le minacce alla sicurezza pubblica e la prevenzione delle stesse», art. 3(46).

[18] E cioè quando ciò sia strettamente necessario per la ricerca di vittime di sottrazione, tratta o sfruttamento sessuale, nonché di persone scomparse; la prevenzione di minacce imminenti e sostanziali alla vita o all'incolumità delle persone, incluse minacce terroristiche; e l'identificazione o localizzazione di sospettati per reati gravi (elencati nell'allegato II), punibili con almeno quattro anni di reclusione (art. 5, lett. h).

[19] V., per esempio, il regolamento (UE) 2017/745 del Parlamento europeo e del Consiglio, del 5 aprile 2017, relativo ai dispositivi medici, il cui considerando n. 89 precisa: «Il presente regolamento rispetta i diritti fondamentali e i principi riconosciuti, in particolare, dalla Carta e, segnatamente, la dignità umana, l'integrità della persona, la protezione dei dati di carattere personale, la libertà delle arti e delle scienze, la libertà d'impresa e il diritto di proprietà. Gli Stati membri dovrebbero applicare il presente regolamento osservando tali diritti e principi». O, ancora, 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 (c.d. GDPR).

valutazione d'impatto sui diritti fondamentali[20]. In tal contesto, la valutazione dei rischi per i diritti umani è equiparata a quella per la salute e la sicurezza, e quanto più elevato è l'impatto negativo di un sistema di IA sui diritti fondamentali tutelati dalla Carta, tanto maggiore sarà il rischio ad esso correlato. L'approccio dell'UE merita di essere certamente elogiato, perché introduce un criterio di conformità qualitativo che va oltre la mera aderenza a requisiti tecnici: gli sviluppatori, infatti, devono anche predisporre misure concrete per mitigare i potenziali effetti dannosi sugli individui e, laddove non sia possibile definire interventi adeguati, sono obbligati a sospendere la diffusione del sistema, informando tempestivamente i fornitori di IA e le autorità nazionali competenti.

Tuttavia, da un attento esame del regolamento, emergono alcune criticità che sollevano interrogativi sull'effettiva capacità dello stesso di garantire una piena tutela dei diritti fondamentali e su cui è necessario, pertanto, soffermarci brevemente.

Anzitutto, l'eccessiva vaghezza e ampiezza delle norme che disciplinano i sistemi di IA classificati come ad alto rischio implicano inevitabilmente il ricorso a norme tecniche armonizzate e a meccanismi di certificazione che ne agevolino l'interpretazione. Questi strumenti, sebbene utili, sono elaborati da organismi di natura privata, che potrebbero non possedere le competenze adeguate a regolare in modo efficace settori così sensibili. Ciò genera, quindi, dubbi circa l'opportunità di affidare a tali soggetti la definizione di parametri normativi che incidono su questioni cruciali per la protezione dei diritti umani, elemento centrale dell'impianto normativo in questione.

In secondo luogo, desta preoccupazione l'eventualità che la valutazione del rischio per i diritti fondamentali – che presenta già la criticità di essere condotta da soggetti che potrebbero giungere a stime disomogenee o non essere in grado di compiere le complesse operazioni di bilanciamento alla base di questo processo – possa seguire gli stessi criteri e metodi di valutazione del rischio utilizzati per la sicurezza dei prodotti. Invero, mentre il pericolo di insicurezza dei prodotti può essere quantificato mediante calcoli computazionali, per i diritti umani non è possibile fare affidamento su dati numerici o modelli matematici: l'applicazione di un approccio tecnico, quindi, potrebbe comportare il rischio di una gestione superficiale e negligente nella protezione di tali diritti.

Tuttavia, non è ancora possibile affermare con certezza se questo strumento si rivelerà inefficace, né se, al contrario, l'approccio europeo alla regolazione dell'IA riuscirà ad affermarsi come standard riconosciuto a livello globale, visto l'intento del legislatore UE di conferirgli un effetto extraterritoriale. Il successo del nuovo quadro normativo dipenderà in larga misura dall'efficacia dei controlli esercitati dalle autorità competenti e dalle istituzioni dell'Unione[21], in particolare dalla Commissione e dalla Corte di giustizia, cui spetta rispettivamente il compito di garantire la corretta applicazione e interpretazione della normativa.

Aprile 2025

[20] *AI Act*, art. 27.

[21] In argomento, v. S. VILLANI, *Il sistema di vigilanza sull'applicazione dell'AI Act: ognun per sé?*, in *Rivista Quaderni AISDUE*, fascicolo speciale n. 2/2024, *La nuova disciplina UE sull'intelligenza artificiale*, p. 125 ss.