RIVISTA DI
DIRITTI COMPARATI

Self-Regulation, Regulation e "contro-regolamentazione". Le nuove tendenze: diritto, diritti e intelligenza artificiale *

Costanza Nardocci

SOMMARIO: 1. *Si sta come d'autunno sugli alberi le foglie*… – 2. L'era delle (non) convergenze: gli approcci, i contenuti. – 3. Un primo approdo: l'Europa e il suo sì alla regolamentazione delle tecnologie di intelligenza artificiale. – 3.1. L'*AI Act*: rischi da monitorare, ma diritti senza *accountability*? – 3.2. Sulla scia dell'Unione Europea? Il Consiglio d'Europa e il primo trattato globale in tema di intelligenza artificiale. – 4. Un secondo. Il "no" degli Stati Uniti e la nuova era della contro-regolamentazione. – 4.1. Prima del "no": l'approccio olistico dell'amministrazione federale precedente. – 4.2. E le scelte di campo statali: da e oltre il Colorado. – 5. Un caso a sé: una "contro-regolamentazione" alternativa: il Canada e la *AI-countergovernance*. – 6. Un terzo. La via di mezzo di chi, dopo l'Unione Europea, vorrebbe regolamentare: il Brasile, la Corea del Sud, la Cina. – 6.1. Il Regno Unito, che parte dai "principi". – 7. Un quarto. La dimensione sovranazionale: le altre organizzazioni internazionali regionali, le Nazioni Unite e la preferenza per strumenti di *soft law*. – 8. Riflessioni conclusive: *si sta* (ancora) tra assonanze e dissonanze.

1. Si sta come d'autunno sugli alberi le foglie…

Il 2025 si è aperto con alcune novità importanti sul versante dei rapporti tra regolamentazione e intelligenza artificiale.

La principale risiede nella presa di posizione netta della nuova amministrazione statunitense, che ha ribaltato i timidi tentativi della precedente di avvicinarsi al continente europeo, offrendo una propria proposta alternativa – ispirata all'esigenza di assoggettare le nuove tecnologie di intelligenza artificiale ad una qualche forma di controllo, anche se più lasca – al Regolamento approvato dalle istituzioni dell'Unione Europea nel 2024, riducendo i margini di azione del c.d. *Brusselles effect*[1].

Una novità da sottolineare, perché espressiva di una linea di tendenza sinora mai così esplicita: una sorta di opposizione vera e propria nei confronti dell'ingresso del diritto e delle sue regole nel settore dell'innovazione tecnologica. Non si tratta di una interpretazione dei primi mesi del 2025, ma del prodotto di alcuni atti che hanno segnato le tappe dei nuovi rapporti tra continenti, europeo e nordamericano in particolare.

Si fa riferimento, per primo, all'esito del *Summit* di Parigi, svoltosi nel febbraio del 2025, e concluso con un accordo siglato da tutte le delegazioni degli Stati partecipanti con le sole eccezioni di Stati Uniti, appunto, del Regno Unito. Il secondo

---

* L'articolo è stato sottoposto, in conformità al regolamento della Rivista, a *double-blind peer review*.

[1] In letteratura, ha impiegato l'espressione *Brusselles effect* per descrivere la capacità del diritto dell'Unione Europea di influenzare altri ordinamenti giuridici A. Bradford, *The Brussels Effect*, in *Northwestern University Law Review*, 2012, 1 e ss.

atto o dato a supporto della tesi della qui denominata "contro-regolamentazione" è rappresentato dall'annullamento, da parte della nuova amministrazione, dell'*Executive Order* adottato dall'amministrazione *Biden-Harris* che aveva costituito il punto più estremo raggiunto dalla strategia federale in tema di intelligenza artificiale.

Tra i due estremi – continente europeo, da un lato, con Unione Europea e Consiglio d'Europa, e gli Stati Uniti, dall'altro – si collocano almeno altre due modalità di affrontare le relazioni tra intelligenza artificiale e diritto. Numerosi sono, infatti, gli esempi di ordinamenti giuridici, che hanno proposte di regolamentazione in corso di approvazione o, almeno, all'esame dei rispettivi organismi rappresentativi, come il Brasile, il Regno Unito.

A questo gruppo di Stati si affiancava, almeno fino alla fine del 2024, anche il Canada, dove la proposta di legge è stata, però, stralciata dal dibattito parlamentare. Il caso canadese è emblematico, discostandosi dal nuovo approccio statunitense, "contro" la regolamentazione dell'intelligenza artificiale, da quello europeo, "per" la regolamentazione, infine, da quello accolto da quegli ordinamenti che sono in attesa di approvare le rispettive prime legislazioni organiche in materia. Se ne distanzia, in quanto il fallimento della proposta di legge costituisce il prodotto dell'omesso coinvolgimento nella discussione parlamentare di *stakeholders* rappresentativi di categorie a rischio di subire pregiudizi dal ricorso alle tecnologie di intelligenza artificiale, secondo una tendenza che la letteratura definisce *AI-Countergovernance*. Si contesta, cioè, non la regolamentazione di per sé, ma che quest'ultima non consideri adeguatamente i diritti e gli interessi di tutti i soggetti coinvolti, minoranze e soggetti svantaggiati per primi.

In una via di mezzo, infine, si collocano, e con la sola eccezione del Consiglio d'Europa, le organizzazioni internazionali che, a partire dalle Nazioni Unite sino a quelle regionali, hanno ripiegato su strumenti di *soft law* piuttosto che seguire il tracciato segnato dalla Convenzione quadro di stampo europeo.

Il saggio mira a considerare ciascuna di queste impostazioni allo scopo di porre, nei limiti dell'inarrestabile dinamismo attuale, qualche punto fermo che consenta di avanzare alcune ipotesi prognostiche rispetto al futuro prossimo dei rapporti tra persona (e i suoi diritti) e tecnologie di intelligenza artificiale[2].

In definitiva, si tratterà di appurare, anche in attesa di snodi decisivi nel percorso di implementazione dell'*AI Act*, se il 2025 porterà ad un capovolgimento dei risultati del biennio 2023-2025 oppure se si sta assistendo ad una battuta di arresto,

[2] Approfondiscono le relazioni tra intelligenza artificiale e diritti, nella prospettiva del diritto costituzionale, A. Simoncini, *L'algoritmo incostituzionale: intelligenza artificiale e il futuro delle libertà*, in *BioLaw Journal*, 2019, 63 e ss.; B. Caravita, *Principi costituzionali e intelligenza artificiale*, in N. Rotundo (a cura di) *Intelligenza artificiale: il diritto, i diritti, l'etica*, Giuffrè, 2020, 451 e ss.; A. D'Aloia, *Il diritto verso "il mondo nuovo". Le sfide dell'Intelligenza Artificiale*, in *BioLaw Journal*, 2019, 3 e ss.

temporanea, che però contraddice le convergenze che sul finire del 2024 sembrava si stessero affacciando su scala globale[3].

2.L'era delle (non) convergenze: gli approcci, i contenuti

Si è premesso che il 2025 si apre con qualche disarmonia che, viceversa, il biennio 2023-2024 sembrava aver, viceversa, indirizzato verso una loro parziale ricomposizione.

Si tratta di aspetto che non stupisce. La storia dell'intelligenza artificiale nei suoi rapporti con il diritto, e con il diritto costituzionale e sovranazionale in particolare, è stata, infatti, sin dai suoi albori segnata da complessità che hanno investito, in estrema sintesi, due ambiti o prospettive principali.

La prima ha investito gli approcci, diametralmente opposti, quanto alla necessità ed opportunità o meno di normare le nuove tecnologie di intelligenza artificiale tramite il ricorso a sistemi più o meno stringenti di norme. Si tratta della nota contrapposizione tra non-regolamentazione (*self-regulation*) e regolamentazione, che ha contraddistinto la prima fase della storia delle relazioni tra diritto e intelligenza artificiale e che si è concretizzata nella prima, anche se blanda, opposizione Unione Europea, da un lato, Stati Uniti, dall'altro.

La seconda, e laddove si è, invece, optato per la regolamentazione, riguarda i contenuti delle medesime. Eterogenee sono state, così, le priorità assegnate alle esigenze di tutela, cioè alle posizioni giuridiche soggettive, da rendere oggetto degli strumenti di salvaguardia. In qualche caso, almeno inizialmente in quasi tutti, la protezione dei dati e la *privacy* hanno occupato una posizione preponderante e di assoluto primo piano, quasi offuscando altri diritti, di pari rilievo costituzionale e sovranazionale, che hanno conquistato solo successivamente un proprio spazio, lasciandone, tuttavia, indietro ancora qualcuno. Il riferimento, tra tutti, e nonostante le dichiarazioni di intenti, ai principi di eguaglianza e di non discriminazione.

Entrambe le prospettive richiamate hanno conosciuto andamenti alterni nel panorama globale e nel corso degli ultimi anni, ma sono i primi mesi del 2025 che sembrano testimoniare un vero e proprio momento di rottura.

Sul piano del metodo (la prima delle due prospettive) regolamentazione *versus* non-regolamentazione, il 2025 ha portato alla ribalta una tendenza ancora diversa: la c.d. contro-regolamentazione. Non si tratta, più, cioè, come ai tempi della *self-regulation*, di una preferenza per un'impostazione astensionistica che evoca le teorie della mano

[3] Offre una sintetica, ma efficace, analisi della situazione in cui versa l'Unione Europea dinanzi alla nuova impostazione statunitense in tema di intelligenza artificiale, H. Ruschemeier, *The De-Regulatory Turn of the EU Commission*, in *Verfassungsblog*, 2025, consultabile al seguente Link: https://verfassungsblog.de/the-de-regulatory-turn-of-the-eu-commission/.

invisibile di Smithiana memoria, quanto piuttosto di una avversione nei confronti della positivizzazione di norme a cui assoggettare le tecnologie di intelligenza artificiale.

La tensione è evidente tra Unione Europea e Stati Uniti e non si è, però, al momento estesa ad altri ordinamenti giuridici nazionali, così come ad organizzazioni internazionali sovranazionali.

Numerosi sono, così, gli esempi di nuove proposte di regolamentazione – un esempio è rappresentato dal Regno Unito – discusse nelle rispettive sedi parlamentari, ancorché con qualche eccezione e non di poco peso. Il riferimento è al Canada che, da primo tra gli ordinamenti giuridici *pro*-regolamentazione, è passato direttamente all'ultimo posto a seguito di un imprevisto stralcio parlamentare della propria proposta legislativa.

A voler tracciare qualche linea direttrice, dinanzi all'intelligenza artificiale, il diritto è stato posto, originariamente, in una posizione di secondo piano (*self-regulation*), a preservare le capacità di sviluppo indipendente del progresso tecnico-scientifico, per imporsi, poi, protagonista imprescindibile del dibattito pubblico – in modo particolarmente evidente nel corso del biennio 2023-2024 –, per divenire, da ultimo, addirittura avversato, contrastato, specie laddove impiegato per tracciare dei confini all'impiego incontrollato delle nuove tecnologie.

Venendo, invece, alla seconda prospettiva, cioè, alla sostanza e ai contenuti delle regolamentazioni, potenziali, proposte ed adottate, si è registrato, analogamente, un certo dinamismo contraddistinto da andamenti convergenti, spinte centrifughe, soluzioni non sovrapponibili.

Vale premettere che, oltre l'oggetto della eterogeneità contenutistica, un aspetto interessante è anche che le sintonie o, meglio, le asincronie tra le proposte non si sono mosse di pari passo con le impostazioni prescelte per normare l'impiego delle tecnologie di intelligenza artificiale.

Non sempre, quindi, ad una tendenza regolatoria ispirata alla logica di una legislazione a maglie larghe si è accompagnata un'impostazione poco sensibile alle esigenze di tutela dei diritti fondamentali. All'opposto, non necessariamente, il dato empirico ed esperienziale, che ci consegnano questi ultimi anni, attesta che alla volontà di assoggettare l'intelligenza artificiale ad un sistema articolato e stringente di norme si affianchi una garanzia ad ampio spettro di tutti i diritti potenzialmente incisi dal funzionamento delle tecnologie in esame. Un elemento che si coglie anche se si confrontano i tentativi regolatori, domestici e sovranazionali, intendendo con questi ultimi gli strumenti vagliati dal diritto internazionale dei diritti umani: minore la precettività dello strumento, come nel caso di documenti di *soft law* oppure dei trattati, maggiore e più evidente il *favor* verso un'impostazione *human-rights based*. Di contro, e guardando invece alle soluzioni vagliate a livello domestico, la tendenza è inversa: ad una maggiore precettività della fonte del diritto preposta a normare il ricorso alle tecnologie di intelligenza artificiale – si pensi all'*AI Act* –, meno intensa si è rivelata la protezione accordata ai diritti fondamentali.

Accanto a questo elemento di incoerenza tra strumento regolatorio e contenuti, nel senso che, come detto, non necessariamente l'effettività, intesa come efficacia applicativa della fonte del diritto, si è rivelata garanzia di altrettanto "pesante" salvaguardia dei diritti, l'opzione per la positivizzazione di norme in tema di intelligenza artificiale, sia essa effettiva o *in itinere*, si è caratterizzata per movimenti e direzioni diversificate.

Qualche esempio.

Vi è chi si è occupato di rischi ed annesse classificazioni delle tecnologie di intelligenza artificiale, vietandone alcune ed assoggettandone altre a controlli in corso di implementazione, l'Unione Europea. Chi, prima dei rischi, aveva preso le mosse dalle esigenze di tutela dei diritti, superando l'esclusività della dimensione della *data protection*, gli Stati Uniti dell'amministrazione *Biden-Harris*. Altri, che hanno deciso di abbandonare approcci di chiusura all'impiego dei sistemi di intelligenza artificiale, accingendosi ad adottare normative *AI-friendly* senza divieti o sanzioni, il Giappone[4]. Vi sono, anche, Stati che, pure seguendo l'impostazione in base al criterio del rischio, hanno preferito legare il primo ai profili di responsabilità derivanti dal malfunzionamento dei sistemi di intelligenza artificiale, il Brasile, oppure, e restando in Sud America, che, pure seguendo da vicino l'impostazione europea, hanno optato per una maggiore enfasi sulle strategie di tutela dei diritti umani, il come il Cile[5]. Ancora, chi, pure non proponendo una categorizzazione delle tecnologie sul modello dell'*AI Act*, ha optato per una regolamentazione puntuale da riservare alle sole suscettibili di

[4] Il Giappone non ha, allo stato attuale, ancora alcuna legislazione organica in materia di intelligenza artificiale, a testimonianza del *favor* a lungo mantenuto verso un approccio morbido nei confronti della disciplina del funzionamento delle tecnologie di intelligenza artificiale, che, per lungo tempo, ha contraddistinto l'impostazione di *self-regulation* che continua a identificare l'ordinamento giuridico giapponese. Tuttavia, anche sulla scia della approvazione dell'*AI Act* da parte dell'Unione Europea, negli ultimi anni si registra una lieve inversione di tendenza. Nel 2024 è stata, così, presentata una prima proposta di regolamentazione denominata *the Basic Act on the Advancement of Responsible AI*, che si fa portavoce, per la prima volta entro l'ordinamento giuridico giapponese, di una impostazione di *hard law*. Merita, inoltre, dare conto della pubblicazione, ad inizio febbraio 2025, dell'*Interim Report*, redatto dall'*AI Policy Study Group*, dal quale traspare la preferenza per un approccio morbido alla regolamentazione che si rimetterebbe direttamente alla volontaria adesione delle aziende alle norme adottate a livello nazionale dallo Stato.

[5] Il Cile, nel dicembre 2021, ha approvato la *National Intelligence Policy*, che pone attualmente l'ordinamento cileno tra gli Stati sud-americani più avanzati in materia di intelligenza artificiale. Non è, però, il solo. Altre iniziative da segnalare si riscontrare, tra le molte: l'Argentina, che, nel 2023, ha visto la presentazione in seno al Congresso nazionale di numerose proposte di legge finalizzate alla regolamentazione delle tecnologie di intelligenza artificiale, senza che attualmente sia possibile pronosticare se e quale proposta conoscerà, poi, effettiva approvazione; la Colombia che, nel 2019, ha introdotto una propria strategia nazionale in tema di IA con la *National Policy for Digital Transformation and Artificial Intelligence*, e, nel 2020, ha istituito una *Task Force*, deputata ad occuparsi dello sviluppo e della implementazione dei sistemi di intelligenza artificiale; soprattutto, il Perù con la sua *Ley que promueve el uso de la inteligencia artificial en favor del desarrollo económico y social del país*, n. 31814 del 2023, vera e propria regolamentazione in tema di intelligenza artificiale.

esporre a pregiudizio la sicurezza nazionale, la Cina. Un panorama, infine, dal quale non va escluso il contributo del Sud del Mondo, esposto al rischio di una nuova era di colonialismo, tecnologico questa volta, ma che vede, però, al momento una leadership indiscussa dell'India[6].

Sul versante dei contenuti, la dimensione sovranazionale ha, infine, portato un contributo importante. Il Consiglio d'Europa, per parte sua e come si dirà, delineando in maniera precisa i diritti *AI-based* e le Nazioni Unite, appuntando l'attenzione sulle esigenze di inclusione e di cooperazione tra Nord e Sud del mondo[7], che dovranno sempre più caratterizzare gli approcci regolatori in materia di intelligenza artificiale[8].

Un panorama, in definitiva, complesso sia sotto il versante del "come" il diritto ha e sta approcciando e tentando di governare le nuove tecnologie, sia in relazione ai contenuti, e con essi, ai diritti, a cui gli interventi regolatori hanno scelto di assegnare priorità.

3. *Un primo approdo: l'Europa e il suo sì alla regolamentazione delle tecnologie di intelligenza artificiale*

L'Unione Europea ha rappresentato per lungo tempo il caso più emblematico di *favor* per la regolamentazione delle tecnologie di intelligenza artificiale.

Il dibattito, concretizzatosi nella bozza pubblicata nell'aprile del 2021[9] ha, poi, come noto, avviato un intenso dibattito parlamentare che ha attraversato gli anni che

---

[6] In proposito, merita segnalare che, a partire dal lancio dell'*India AI Mission* nel 2024, l'ordinamento giuridico indiano ha conquistato una posizione di sicuro primo piano quale leader globale nella ricerca e nello sviluppo delle tecnologie di intelligenza artificiale. In questa prospettiva, il Governo ha, inoltre, lanciato ad inizio 2025 il suo *Indigenous Affordable, Secure AI Model*. L'iniziativa mira a rendere i sistemi di intelligenza artificiale accessibili e vantaggiose in favore di una vasta gamma di utenti, ricercatori, sviluppatori, studenti, imprese. La caratteristica fondamentale del modello è l'uso delle lingue indiane, che consentirà al modello di rispondere alla diversità linguistica e culturale del Paese e di essere personalizzato per applicazioni locali e regionali. Un modello, in definitiva, che risponde, almeno nelle intenzioni, ad una forte logica di inclusione e di non discriminazione.

[7] In Africa, esempi di ordinamenti giuridici che hanno iniziato ad impiegare le normative esistenti in materia di protezione dei dati per assicurare il governo delle tecnologie di intelligenza artificiale. Si vedano, tra questi, gli esempi di Nigeria e Kenya.

[8] Sulla necessaria inclusione degli Stati del Sud del Mondo nella discussione globale sulle strategie di regolamentazione dell'intelligenza artificiale si è discusso in occasione del recente Summit di Parigi 2025. Si veda, in tema, quanto riportato a commento dei contenuti del Summit da parte del quotidiano *LeMonde* al seguente link: https://www.lemonde.fr/en/economy/article/2024/12/10/the-paris-ai-summit-will-act-against-the-risks-and-for-the-benefits-of-ai_6735916_19.html.

[9] Il testo integrale può essere letto al seguente Link: https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence. Per un commento a prima lettura della proposta di Regolamento, si vedano, tra i molti, C. Casonato, B. Marchetti, *Prime osservazioni sulla proposta di regolamento dell'Unione Europea in materia di intelligenza artificiale*, in *BioLaw Journal*, 2021, 415 e ss.; M. Kop, EU *Artificial Intelligence Act: The European*

hanno separato l'approvazione del testo definitivo del 13 giugno 2024[10], entrato in vigore dal primo agosto dello stesso anno.

Non si vuole tornare, in questa sede, sui passaggi che hanno contraddistinto la successione di emendamenti riflesso dell'avvicendarsi delle diverse Presidenze, bensì delineare i punti fermi o caratterizzanti l'impostazione dell'Unione Europea all'intelligenza artificiale.

Seguendo la duplice lettura – approccio, da un lato, contenuti/sostanza, dall'altro – descritta nel paragrafo che precede, il Regolamento dell'Unione Europea, c.d. *AI Act*, presenta due caratteristiche essenziali.

Quanto alla prima e, cioè all'approccio, da segnalare non è solo la volontà regolatoria, ma, soprattutto, la preferenza per un sistema di norme robusto, fatto di regole rigide e prodotto di scelte calate dall'alto dal legislatore dell'Unione. Scelta da sottolineare se non altro perché si rivela non in tutto coerente con l'oggetto del Regolamento. Le tecnologie di intelligenza artificiale sono di natura soggette ad evoluzioni repentine, circostanza che avrebbe dovuto suggerire una maggiore flessibilizzazione delle norme introdotte alla luce del principio della c.d. neutralità tecnologica. Detto altrimenti, l'ontologia normativa dell'*AI Act* poco si concilia con quella, viceversa, mutevole dell'innovazione scientifica e tecnologica, che esigerebbe norme elastiche dotate di un qualche grado di ultrattività tale da adeguare l'impianto normativo alla prima.

La seconda caratteristica riguarda, invece, i contenuti del Regolamento, di cui si intende porre in rilievo i profili di maggiore tipicità, quelli, cioè, che meglio caratterizzano l'esperienza regolatoria continentale-europea.

Un primo, definitorio, attiene alla nozione di intelligenza artificiale[11].

Approach to AI, in *Transatlantic Antitrust and IPR Developments*, 2021; W.G. Voss, *AI Act: the European Union's Proposed Framework Regulation for Artificial Intelligence Governance*, in *Journal of Internet Law*, 2021, 7 e ss.; M. Veale, F.Z. Borgesius, *Demystifying the Draft EU Artificial Intelligence Act*, in *Computer Law Review International*, 2021, 97 e ss.; L. Floridi, *The European Legislation on AI: A Brief Analysis of its Philosophical Approach*, 2021, consultabile al seguente Link: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3873273.

[10] Per un commento a prima lettura del testo definitivo dell'*AI Act* anche tramite uno sguardo ai mesi immediatamente antecedenti la sua approvazione, si rinvia, tra i molti, a C. Casonato, G. Olivato, *AI Regulation in Europe: Exploring the Artificial Intelligence Act*, in A. Fabris, S. Belardinelli (a cura di), *Digital Environments and Human Relations*, Springer, 2024, pp. 87 e ss.; C. Novelli, F. Casolari, A. Rotolo, M. Taddeo, L. Floridi, *Taking AI risks seriously: a new assessment model for the AI Act*, in *AI and Society*, 2024, 2493 e ss.; A. Alaimo, *Il Regolamento sull'Intelligenza Artificiale: dalla proposta della Commissione al testo approvato dal Parlamento. Ha ancora senso il pensiero pessimistico?*, in *Federalismi.it*, 2023, 133 e ss.; C. Novelli, *L'Artificial Intelligence Act (AIA) Europeo: alcune questioni di implementazione*, in *Federalismi.it*, 2024, 95 e ss. Si vedano, inoltre, i contributi pubblicati all'interno del *Forum "Il Regolamento sull'intelligenza artificiale"*, in *Rivista del "Gruppo di Pisa"*, 2023.

[11] La nozione accolta nella versione definitiva del testo di Regolamento non è stata esente da critiche nel tortuoso percorso, che ha condotto all'approvazione dell'*AI Act*. L'art. 3, nella sua prima versione, recitava come segue: "'sistema di intelligenza artificiale' (sistema di IA): un software sviluppato con una o più delle tecniche e degli approcci elencati nell'allegato I, che può, per una determinata serie

L'Unione Europea sceglie di seguire le orme dell'OECD, secondo cui per intelligenza artificiale deve intendersi "un sistema automatizzato progettato per funzionare con livelli di *autonomia* variabili e che può presentare adattabilità dopo la diffusione e 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 [*corsivo nostro*]"[12].

Non ci si sofferma, in questa sede, sulla *ratio* dell'opzione definitoria, espressione di un chiaro *endorsement* per l'impostazione *science-based*, volta a ricondurre sotto l'ombrello della nozione in esame le sole tecnologie che possiedano un grado di autonomia rispetto alla componente umana ed una capacità di funzionamento non necessariamente causalmente collegata alla condotta della persona. Il testo valorizza il sostantivo, "intelligenza", seguendo una lettura che qualifica intelligenza artificiale solo quelle tecnologie in grado di replicare abilità tipicamente umane: per tutte, la capacità di compiere scelte ragionevoli.

La definizione da ultimo accolta dal Regolamento dell'Unione e qui in discussione sicuramente amplia le maglie delle versioni precedenti, rispondendo all'obiezione secondo cui, all'opposto, una troppo ristretta avrebbe sottratto a controlli una vasta serie di *software*. Al tempo stesso, la scelta per una definizione ampia sembra,

---

di obiettivi definiti dall'uomo, generare output quali contenuti, previsioni, raccomandazioni o decisioni che influenzano gli ambienti con cui interagiscono". Su questa prima proposta si sono sviluppate opinioni contrastanti. Da un lato, si collocava la posizione di chi riteneva eccessivamente ridotta la definizione proposta, *rea* di non includere in modo adeguato *software* non *machine learning*. In questo senso, J. Ellul, *Should we regulate Artificial Intelligence or some uses of software*, in *Discover Artificial Intelligence*, 2022; H. Ruschemeier, *AI as a challenge for legal regulation: the scope of application of the artificial intelligence act proposal*, in *ERA Forum*, 2023, 361 e ss.; N.A. Smuha, E. Ahmed-Rengers, A. Harkens, W. Li; J. MacLaren; R. Piselli; K. Yeng, *How the EU Can Achieve Legally Trustworthy AI: A Response to the European Commission's Proposal for an Artificial Intelligence Act*, 2021, disponibile al seguente Link: https://ssrn.com/abstract=3899991 or http://dx.doi.org/10.2139/ssrn.3899991. Dall'altro, vi era chi giudicava troppo ampia la nozione proposta che, avrebbe dovuto essere circoscritta soltanto ad alcune tipologie di sistemi machine learning. Per un approfondimento sul dibattito che ha interessato la nozione di IA, si rinvia, diffusamente, a C.T. Castàn, *The legal concept of artificial intelligence: the debate surrounding the definition of AI systems in the AI Act*, in *Biolaw Journal*, 2024, 303 e ss. Volendo, almeno per la prima fase del dibattito, anche, C. Nardocci, *L'Unione Europea al cospetto dell'intelligenza artificiale: verso la tutela dei diritti fondamentali? Forse*, in *Quaderni Costituzionali*, 2022, 385 e ss., in cui si dà conto criticamente della prima proposta di modifica del testo originario del Regolamento, pubblicato il 21 aprile 2021, adottata dal Consiglio che aveva ridefinito in senso restrittivo della nozione di IA.

[12] Si veda l'art. 3, par. n. 1, dell'*AI Act*. La definizione è stata inizialmente suggerita nel 2019 nell'ambito dei *Principles for trustworthy AI*, poi aggiornati nel 2024. Il testo integrale, nella versione aggiornata, è consultabile al seguente Link: https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449. Ad integrazione della definizione contenuta nel testo del Regolamento, meritano, in questa sede, essere ricordate le linee guida, adottate in materia dalla Commissione dell'Unione Europea. Il riferimento è agli *Orientamenti sulle pratiche vietate in materia di intelligenza artificiale (IA)*, pubblicati il 4 febbraio 2025 e consultabili al seguente Link: https://digital-strategy.ec.europa.eu/en/library/commission-publishes-guidelines-prohibited-artificial-intelligence-ai-practices-defined-ai-act.

poi, essere stata ridimensionata dalla proposta categorizzazione delle tecnologie di intelligenza artificiale e, in particolare, della norma che vuole sottratte alla c.d. valutazione di impatto, di cui agli artt. 9 e 27 del testo, tutte quelle che non siano né vietate, né ad alto rischio.

Qui si giunge al secondo aspetto cruciale dell'*AI Act*, cioè la classificazione in base al rischio delle tecnologie di intelligenza artificiale[13], a cui si affiancano le annesse conseguenze sul piano degli obblighi posti in capo agli Stati membri[14].

La scelta di elevare il rischio a criterio di distinzione tra tipologie di intelligenza artificiale non costituisce tratto esclusivo dell'azione regolamentare dell'Unione Europea. Piuttosto, ciò che contraddistingue il Regolamento dell'Unione è la decisione di vietarne alcune, salvo alcune e invero non irrilevanti né rare eccezioni; la soggezione solo di quelle considerate "ad alto rischio" ad uno scrutinio in punto di loro potenziale lesività sul piano dei diritti individuali; l'esclusione da qualsiasi controllo per tutte le altre. In altre parole, la peculiarità risiede nelle conseguenze più che nell'elemento impiegato per distinguere una categoria dall'altra e, cioè, controlli – e, quindi, potenziali sanzioni – solo per le prime due categorie, libertà di utilizzo per la terza

Una precisazione è utile per comprendere le ricadute concrete di un simile approccio.

Il Regolamento individua e dettaglia le sole tecnologie incluse nelle prime due categorie – vietate e ad alto rischio –, mentre ripiega su un criterio residuale quanto a quelle, viceversa, riconducibili alla terza categoria. Nessuna obiezione fin qui, verrebbe da osservare. Tuttavia, la precisazione è necessaria per due ragioni: la prima ha a che fare con la rigidità della categorizzazione proposta, che non ammette movimenti inter-categorie al mutare ed evolversi del progresso scientifico e tecnologico; la seconda, collegata alla prima, riguarda, poi, gli effetti dell'omessa verifica in termini di compliance delle tecnologie dì intelligenza artificiale ritenute, ad oggi, non foriere di pericoli sul piano dei diritti fondamentali.

Con riferimento al primo punto, ossia alla tipizzazione delle tecnologie inquadrabili nelle prime due categorie, se ne potrebbe criticare, anzitutto, la immutabilità. La tipizzazione che si accompagna a scelte normative fisse stride con il

---

[13] Sulla riferibilità della nozione di rischio alle tecnologie di intelligenza artificiale, si vedano F.L.G. Faroldi, *Risk and artificial general intelligence*, in *AI & Society*, 2024, 1 e ss.; C. Orwat, J. Bareis, A. Folberth et Al., *Normative Challenges of Risk Regulation of Artificial Intelligence*, in *Nanoethics*, 2024, pp. 1 e ss.

[14] Un approfondimento su tutte le declinazioni che l'approccio basato sul rischio conosce all'interno del testo definitivamente approvato del Regolamento dell'Unione Europea in materia di intelligenza artificiale è offerto da P. Voigt, N. Hullen, *The EU AI Act. Answers to Frequently Asked Questions*, Springer Nature, 2024; F.L.G. Faroldi, *Risk and artificial general intelligence*, in *AI and Society*, 2024, consultabile al seguente Link: https://doi.org/10.1007/s00146-024-02004-z; C. Orwat, J. Bareis, A. Folberth, et Al., *Normative Challenges of Risk Regulation of Artificial Intelligence*, in *Nanoethics*, 2024, 1 e ss. Tra i molti, approfondisce l'utilizzo del criterio del rischio nel quadro del diritto dell'Unione Europea M.E. Gonçalves, *The risk-based approach under the new EU data protection regulation: a critical perspective*, in *Journal of Risk Research, Taylor & Francis Journals*, 2020, 139 e ss.

dinamismo dell'innovazione tecnologica e rischia di consentire l'utilizzo di nuovi sistemi altrettanto (potenzialmente) nocivi per il solo fatto di non essere stati previamente oggetto dell'intervento del legislatore dell'Unione.

Discorso analogo vale anche per la residualità. Le tecnologie considerate non ad alto rischio e, ovviamente, quelle non vietate, vanno esenti dalla valutazione di impatto di cui all'art. 27 del Regolamento che ha ad oggetto, invece, le sole ad alto rischio, con l'effetto che si potrà appurare solo *ex post*, a violazione e danno verificatisi, la pericolosità di una tecnologia di intelligenza artificiale apparentemente innocua o, comunque, non inserita nei due elenchi di cui agli art. 5, tecnologie vietate, art. 6, ad alto rischio[15].

Terzo elemento da richiamare del Regolamento, anche qui da circoscrivere, è la discussa valutazione di impatto di cui riferisce l'art. 27 dell'*AI Act* e, attualmente, al centro di un intenso dibattito in seno alla Commissione per declinarne le concrete modalità di realizzazione[16].

La circostanza secondo cui le tecnologie di intelligenza debbano essere in qualche misura controllate e non possano tradursi in violazioni di diritti fondamentali è dato incontestabile, almeno per chi ha optato per la strada della regolamentazione. Ciò che è, invece, meno scontata è la delimitazione dell'oggetto del controllo, che il Regolamento riserva alle sole tecnologie ritenute ad alto rischio, replicandosi, qui, le criticità sopra descritte quanto alla predeterminazione rigida dei sistemi da considerare vietati. Una rigidità che, ancora una volta, sarebbe preferibile temperare con qualche elemento di flessibilità.

Accanto all'oggetto della valutazione di impatto, cioè ammessa soltanto per quelle tecnologie di intelligenza artificiale che il Regolamento sceglie di qualificare come ad alto rischio, un secondo aspetto concerne i tempi. La norma di cui all'art. 27 stabilisce che la valutazione di risk *assesment* debba svolgersi prima della messa in funzione definitiva della tecnologia di intelligenza artificiale.

Nessuna obiezione su questo punto, che, peraltro, non separa, ma, anzi, accomuna la scelta europea con quelle di coevi tentativi regolatori. Piuttosto, resta

---

[15] In proposito, merita, peraltro, precisare che la delimitazione delle tecnologie di intelligenza artificiale assoggettate alla procedura di *risk assesment*, prefigurata ai sensi dell'art. 27 *AI Act*, risulta essere ulteriormente circoscritta non coprendo tutte le categorie "ad alto rischio" ricavabili dalla lettura dell'art. 6, bensì riferendosi soltanto a quelle elencate ai sensi dell'Allegato III al Regolamento UE.

[16] Diritti fondamentali che, nell'impianto del Regolamento dell'Unione Europea, sono i seguenti: "il diritto alla dignità umana, il rispetto della vita privata e della vita familiare, la protezione dei dati personali, la libertà di espressione e di informazione, la libertà di riunione e di associazione e il diritto alla non discriminazione, il diritto all'istruzione, la protezione dei consumatori, i diritti dei lavoratori, i diritti delle persone con disabilità, l'uguaglianza di genere, i diritti di proprietà intellettuale, il diritto a un ricorso effettivo e a un giudice imparziale, i diritti della difesa e la presunzione di innocenza e il diritto a una buona amministrazione", così il par. n. 28 del *Considerando*. Si sofferma sulle principali novità del Regolamento dell'Unione Europea con particolare attenzione alle sue ricadute sul grado di protettività offerto ai diritti fondamentali F. Donati, *La protezione dei diritti fondamentali nel Regolamento sull'intelligenza artificiale*, in *Rivista AIC*, 2025, 1 e ss.

aperto l'interrogatorio di quale sia oppure debba essere lo spazio da riservare a monitoraggi che consentano una valutazione anche *ex post* e non solo *ex ante* del funzionamento della tecnologia sottoposta a controllo. Il Regolamento europeo non contiene prese di posizione sul tema e sarà da verificare come, eventualmente, deciderà di porsi il *Code of Practice* di prossima pubblicazione. Non si può, in ogni caso, omettere di rilevare l'opportunità di una verifica quanto più possibile costante dei meccanismi di funzionamento dei sistemi di intelligenza artificiale, tanto più opportuna in considerazione della estrema varietà dei settori che ne faranno uso e delle finalità a cui sarà preposto il loro impiego. Su tutti questi punti, il panorama comparato, in particolare alcune leggi statali statunitensi, costituiscono esempi meritori a cui guardare.

Il quarto, ma non ultimo, tratto distintivo del Regolamento adottato dalle istituzioni dell'Unione Europea attiene alle modalità con cui il testo pone in relazione le tecnologie di intelligenza artificiale con i diritti fondamentali. Si tratta, cioè, del contenuto e non certo pervasivo approccio *human rights-based*, che, nonostante le critiche mosse alle bozze del Regolamento che si sono succedute dal 2021 in avanti, non ha mai superato del tutto quel modo di intendere l'intelligenza artificiale quale fenomeno intrinsecamente e, per certi aspetti, inscindibilmente legato alla protezione dei dati personali. L'allargamento agli altri diritti e, tra questi, l'eguaglianza e la non discriminazione, si è realizzato, infatti, solo in parte e traspare più dalle dichiarazioni di principio del Preambolo che apre il Regolamento che non dalle norme che ne compongono l'articolato.

Si veda, in questo senso, e pare fare qualche esempio, il Considerando che include tra gli obiettivi del Regolamento l'esigenza di: "garantire un livello costante ed elevato di tutela degli interessi pubblici in materia di salute, sicurezza e diritti fondamentali"[17] dotarsi di "un quadro giuridico […] che istituisca regole armonizzate in materia di intelligenza artificiale […], garantendo […] un elevato livello di protezione degli interessi pubblici, quali la salute e la sicurezza e la protezione dei diritti fondamentali"[18]; assicurare la massimizzazione dei benefici derivanti dalle tecnologie di IA; operare una categorizzazione dei sistemi di IA che poggi sulla valutazione "dell'impatto negativo del sistema di IA sui diritti fondamentali"[19].

Emblematico di questo approccio è, anche, il citato art. 27, che, come detto, richiede di procedere ad uno scrutinio sulla potenziale lesività dei sistemi di intelligenza artificiale ad alto rischio, come definiti a norma dell'art. 6 e dell'*Annex* III, nei confronti dei diritti fondamentali protetti dalla Carta dei diritti fondamentali dell'UE.

Espressivo, infine, della scarsa adesione alla impostazione *human-rights based* è la pressoché assente menzione delle implicazioni dell'intelligenza artificiale sul piano dei principi di eguaglianza e di non discriminazione. Un dato, che si impone ad una prima

[17] Così il punto n. 5 del Considerando.
[18] *Ibidem.*
[19] Così il punto n. 28 del Considerando.

lettura del testo, anzitutto, a motivo dell'omesso riferimento alle Direttive dell'Unione Europea degli anni 2000 in tema di diritto anti-discriminatorio. Se la versione definitiva del testo ha perlomeno introdotto qualche riferimento alle esigenze che l'impiego delle tecnologie di intelligenza artificiale si conformi al rispetto dell'eguaglianza e della non discriminazione – sopperendo alle ancora più evidenti carenze della bozza dell'aprile 2021 –, rimangono aperti due interrogativi, a cui in questa sede si accenna soltanto. Il primo attiene alla assoggettabilità o meno della discriminazione, che scaturisce dal ricorso alle tecnologie di intelligenza artificiale alle norme del diritto anti-discriminatorio classico, su cui tace il Regolamento. Il secondo, invece, è se le tesi avanzate in lettura sulla eterogeneità della discriminazione per come plasmata dall'*agere* della componente automatizzata abbiano o meno trovato cittadinanza nel quadro dell'*AI Act*, con tutto quanto ne consegue in punto di inadeguatezza o scarsa efficacia delle categorie classiche del diritto antidiscriminatorio alla definita discriminazione *AI-based*.

Due interrogativi, che rimangono aperti e su cui si attende una presa di posizione del legislatore dell'Unione, in occasione della pubblicazione del prossimo *Code of Practice* anche alla luce di importanti progetti, che sempre l'Unione Europea ha intrapreso in collaborazione, tra gli altri, con il Consiglio d'Europa, con la finalità di colmare le lacune esistenti e di irrobustire le garanzie di tutela avverso la discriminazione derivante dalle tecnologie di intelligenza artificiale[20].

3.1.L'AI Act: rischi da monitorare, ma diritti senza accountability?

Vi è un ulteriore aspetto, più controverso, ma meno discusso che contraddistingue l'impostazione delle istituzioni dell'Unione Europea quanto alla regolamentazione delle tecnologie di intelligenza artificiale.

Ci si riferisce all'assenza, nel testo del Regolamento, di meccanismi di accertamento della *liability* e, in definitiva, di *enforceability* concreta dell'impianto normativo, soprattutto sul piano della salvaguardia dei diritti fondamentali, che vadano oltre la previsione di cui all'art. 77 dell'*AI Act*, rubricato *Poteri delle autorità che tutelano i diritti fondamentali*. Norma, che dettaglia le obbligazioni delle autorità nazionali,

[20] Il riferimento è al progetto *Upholding equality and non-discrimination by Equality bodies regarding the use of artificial intelligence (AI) in public administrations*. L'*Executive Summary* può essere consultato seguente link: https://www.unia.be/files/20250331_AI-TSI-Executive-summary-gap-analyse-EN.pdf. Il progetto possiede una dimensione geografica delimitata a tre soli Stati membri – Belgio, Finlandia e Portogallo –, rispondendo alle esigenze avanzate dagli *Equality Bodies* rispettivamente operanti nei singoli contesti ordinamentali, ma appare di sicuro rilievo ed esportabile anche entro il, seppure diverso, contesto italiano. Ulteriori informazioni sul progetto possono essere consultate sul sito ufficiale del Consiglio d'Europa al seguente Link: https://www.coe.int/en/web/inclusion-and-antidiscrimination/upholding-equality-and-non-discrimination-by-equality-bodies-regarding-the-use-of-artificial-intelligence-ai-in-public-administrations1.

chiamate ad assolvere a funzioni di supervisione e di controllo sul funzionamento, non lesivo dei diritti umani e non discriminatorio, delle tecnologie ad alto rischio.

Questa lacuna riflette un dibattito che continua ad occuparsi del *come* realizzare la valutazione di impatto, riservata alle tecnologie di intelligenza artificiale ad alto rischio secondo quanto dispone l'art. 27 *AI Act*, e che considera, viceversa, recessivo il tema del *chi* debba rispondere di violazioni dei diritti fondamentali. La distanza tra valutazione dei rischi legati all'impiego delle tecnologie ad alto rischio, da un lato, e accertamento delle responsabilità, dall'altro, stride con le potenzialità e ambizioni applicative del Regolamento, ma è divenuta ormai tratto, purtroppo, caratteristico della *governance* europea dell'intelligenza artificiale.

In questa prospettiva, si inserisce il ritiro, da parte della Commissione dell'Unione Europea, della c.d. *AI Liability Directive*, la *Proposal for a Directive of the European Parliament and of the Council on adapting non-contractual civil liability rules to artificial intelligence*, che, nelle intenzioni dei proponenti, rispondeva all'esigenza di integrare l'impianto normativo tratteggiato dall'*AI Act* allo scopo di apprestare e disciplinare strumenti risarcitori adeguati ed efficaci nell'interesse delle persone e dei loro diritti[21].

Come si legge nel *memorandum* esplicativo, obiettivo della proposta avrebbe dovuto essere "[to ensure] victims of damage caused by AI obtain equivalent protection to victims of damage caused by products in general"[22].

Al tempo stesso, inoltre, la Direttiva mirava a favorire il progressivo contenimento delle incertezze che attualmente circondano l'operato di utilizzatori finali e programmatori di tecnologie di intelligenza artificiale in relazione alle rispettive responsabilità derivanti dal funzionamento di queste ultime.

Tra le novità più significative, la disposizione di cui all'art. 3, rubricato *Divulgazione degli elementi di prova e presunzione relativa di non conformità*, ai sensi del quale si stabiliva che la vittima di un pregiudizio originato da un sistema di intelligenza artificiale avrebbe potuto ottenere dallo Stato membro la messa a disposizione – previa interlocuzione con il fornitore del sistema di intelligenza artificiale – di tutti gli elementi fattuali necessari a comporre l'eventuale quadro probatorio a sostegno della propria

[21] Da richiamare è la Relazione della Commissione dell'Unione Europea, *Relazione sulle implicazioni dell'intelligenza artificiale, dell'Internet delle cose e della robotica in materia di sicurezza e di responsabilità*, al Parlamento, al Consiglio e al Comitato economico e sociale europeo, dove, tra le altre cose e a supporto di un auspicato intervento normativo, si precisava quanto segue: "[s]ebbene l'ottimizzazione delle norme dell'Unione in materia di sicurezza per l'intelligenza artificiale possa contribuire a evitare incidenti, questi possono comunque verificarsi. In tal caso interviene la responsabilità civile. Le norme in materia di responsabilità civile hanno una duplice funzione nella nostra società: da una parte, garantiscono che quanti hanno subito un danno causato da altri ottengano il risarcimento e, dall'altra, creano un incentivo economico per spingere la parte responsabile a evitare di causare il danno. Le norme in materia di responsabilità devono sempre trovare un equilibrio, tutelando i cittadini dai danni e consentendo allo stesso tempo alle imprese di innovare". Il testo integrale della Relazione, pubblicata in data 19 febbraio 2020, può essere consultato al seguente Link: https://eur-lex.europa.eu/legal-content/IT/TXT/HTML/?uri=CELEX:52020DC0064.

[22] Cfr. punto n. 1, sezione rubricata *Context of the proposal.*

istanza. Ispirata al principio della parità delle armi era, poi, la norma successiva che introduceva la regola della presunzione di colpa del convenuto[23], cioè del fornitore del sistema di intelligenza artificiale che si assumeva avesse cagionato il danno, laddove fossero ricorse le condizioni tipizzate dalla disposizione: la dimostrazione, da parte dell'attore, che "la colpa del convenuto o di una persona della cui condotta il convenuto è responsabile consist[a] nella non conformità a un obbligo di diligenza previsto dal diritto dell'Unione o nazionale e direttamente inteso a proteggere dal danno verificatosi"[24]; la ragionevole probabilità "sulla base delle circostanze del caso, che il comportamento colposo abbia influito sull'output prodotto dal sistema di IA o sulla mancata produzione di un output da parte di tale sistema"[25]; la dimostrazione, posta in capo all'attore, "che il danno è stato causato dall'output prodotto dal sistema di IA o dalla mancata produzione di un output da parte di tale sistema"[26].

Nella prospettiva che qui interessa, però, l'aspetto cruciale della proposta di Direttiva risiedeva nella volontà di istituire un meccanismo di responsabilità civile, funzionale a restituire piena efficacia al diritto di accesso al giudice previsto ai sensi dell'art. 47 della Carta di Nizza[27]. In altre parole, nelle intenzioni della proposta di Direttiva e allo scopo di colmare il c.d. *liability* o *responsability gap*[28], chiunque fosse stato leso da un sistema di intelligenza artificiale avrebbe dovuto e potuto essere posto nelle condizioni di ottenere il risarcimento del danno subito, secondo un *corpus* normativo innovativo capace di tenere conto delle specificità delle tecnologie di intelligenza artificiale rispetto alle caratteristiche che seguono pregiudizi causati da prodotti di altra natura.

Il *memorandum* esplicativo risultava particolarmente chiaro ed efficace su questo punto. Si leggeva, così, che "[t]he proposal [would have] enable[d] effective private enforcement of fundamental rights and preserve the right to an effective remedy where AI-specific risks have materialised"[29].

In particolare, e nelle intenzioni dei suoi estensori, la proposta di Direttiva avrebbe dovuto assicurare adeguata protezione avverso violazioni di diritti fondamentali salvaguardati dalla Carta dei diritti fondamentali dell'Unione Europea.

---

[23] Cfr. articolo 4, rubricato *Presunzione relativa del nesso di causalità in caso di colpa*.

[24] Cfr. lett. *a)*.

[25] Cfr. lett. *b)*.

[26] Cfr. lett. *c.)*

[27] Aspetto meritevole di essere segnalato soprattutto poiché le modalità con cui le tecnologie di intelligenza artificiale discriminano sono, spesso, ignote alla vittima (su cui si rinvia, *supra*, la Parte Seconda dello Studio) e la predisposizione di un sistema di norme di diritto positivo che risponda a queste specificità avrebbe rappresentato un punto di svolta importante sia in relazione all'art. 47 che agli artt. 20 e 21 della Carta dei diritti fondamentali dell'Unione Europea, nonché ai corrispettivi principi operanti a livello nazionale.

[28] In tema, si vedano F. Santoni de Sio, G. Mecacci, *Four Responsibility Gaps with Artificial Intelligence: Why they Matter and How to Address them*, in *Philosophy and Technology*, 2021, 1057 e ss.

[29] Cfr. punto n. 1, sotto-sezione rubricata *Fundamental Rights*.

Tra questi: il diritto alla vita e alla integrità fisica, il diritto di proprietà, sino al diritto alla vita privata, l'eguaglianza e la non discriminazione.

A quest'ultimo riguardo – che maggiormente interessa le finalità dello studio –, la proposta di Direttiva non mancava di precisare il proprio rapporto di integrazione e di complementarietà con il diritto antidiscriminatorio dell'Unione Europea, palesando, o meglio, confermando, tra le righe, ma non così tanto, la tesi della non auto-sufficienza delle norme vigenti per il contrasto delle discriminazioni derivanti dalle tecnologie di intelligenza artificiale. Si trattava di un punto estremamente significativo, che avrebbe potuto colmare o spiegare, almeno in via interpretativa, il coevo silenzio dell'*AI Act*.

E, invero, tra le criticità, di cui si è detto, del Regolamento dell'Unione, vi è sicuramente il silenzio quanto al raccordo o (dis)accordo con il diritto antidiscriminatorio vigente, ma soprattutto sui profili di responsabilità, che rimangono scoperti pure di fronte all'obbligazione di procedere alla ormai famosa valutazione di impatto[30].

Sfortunatamente, la proposta di Direttiva non ha trovato l'accordo sperato e nel febbraio del 2025 la Commissione dell'Unione, come premesso, ha optato per il ritiro della bozza[31]. Un passo indietro, che creerà non pochi problemi e sotto un duplice versante, che qui si accenna soltanto. Da un lato, l'assenza dell'auspicato nuovo sistema di verifica e accertamento delle responsabilità legate al funziona-mento delle tecnologie di intelligenza artificiale rischia di rendere la tutela dei diritti fondamentali dei cittadini e delle cittadine dell'Unione sempre più effimera e legata ai contesti dei singoli Stati membri se non, addirittura, delle specificità del caso singolo se portato davanti al giudice.

Da altro angolo prospettico, non sembra si possa dire che vada meglio per chi programma, sviluppa, mette in commercio e utilizza i sistemi di intelligenza artificiale. Assenza di norme in punto di responsabilità non equivale necessariamente ad uno *status quo* a maglie larghe, ma, piuttosto, potrebbe produrre l'effetto opposto con irrigidimenti e divieti che avrebbero quale conseguenza proprio quella imbrigliatura della scienza e della innovazione che non si vorrebbe.

La proposta di Direttiva, insomma, istituiva un delicato equilibrio tra esigenze contrapposte che si troveranno, d'ora in avanti, sempre più in contrapposizione con il timore che a farne le spese saranno, come spesso, i diritti delle persone e dei gruppi e, tra questi, dei e delle più fragili[32], senza che la lacuna lasciata scoperta dalla *Liability*

---

[30] Confermavano la bontà e il pregio della proposta di Direttiva, M. Ziosi, J. Mökander, C. Novelli, F. Casolari, M. Taddeo, L. Floridi, *The EU AI Liability Directive: shifting the burden from proof to evidence*, *SSRN Scholarly Paper. Rochester*, 2003, 1 e ss.

[31] La notizia è stata riportata da numerose testate giornalistiche. In particolare, si veda la notizia pubblicata dalla Reuters e consultabile al seguente Link: https://www.reuters.com/technology/eu-ditches-plans-regulate-tech-patents-ai-liability-online-privacy-2025-02-12/.

[32] Approfondisce le criticità descritte e relative alle carenze sotto il profilo della responsabilità per le violazioni delle norme del Regolamento UE in tema di intelligenza artificiale ma, più in generale, sulle lacune dell'*AI Act* S. Watcher, *Limitations and Loopholes in the EU AI Act and AI Liability Directives:*

Directive possa essere in qualche misura colmata dal *Code of Practice* di prossima adozione[33].

3.2.Sulla scia dell'Unione Europea? La Convenzione Quadro del Consiglio d'Europa

A distanza di quasi due mesi dall'approvazione in via definitiva dell'*AI Act*, il Consiglio d'Europa ha pubblicato, il 17 maggio 2024, la Convenzione Quadro in materia di IA, diritti umani, democrazia e legalità[34].

Frutto del lavoro, avviato nel 2019, dei due Comitati *ad hoc* istituti dal Consiglio d'Europa – il CAHAI[35] e il CAI[36] –, incaricati, il primo, di redigere uno studio di

---

What This Means for the European Union, the United States, and Beyond, in *Yale Journal of Law & Technology*, 2024, 671 e ss.

[33] Il processo di stesura del *General-Purpose AI Code of Practice* è ancora in corso. Lanciato il 30 settembre 2024 dall'*AI Office* con il lancio della plenaria del *Code of Practice*, il lavoro è stato organizzato in quattro *working groups* ed ha portato alla pubblicazione della seconda *Draft* nel dicembre 2024. Si attende la pubblicazione della versione finale del *Code of Practice* per il 2025. Nel momento in cui si scrive, è stata pubblicata la terza bozza del testo. Esprimono criticità sulla effettiva adeguatezza del *Code* a salvaguardare i diritti fondamentali, anzitutto, quelli protetti dalla Carta dei diritti fondamentali dell'Unione Europea un insieme di organizzazioni non governative guidate da *Article19*. Il testo dello *Statement* indirizzato a Henna Virkkunen, Vice-Presidente e commissaria per la sovranità tecnologica, sicurezza e democrazia, può essere letto al seguente Link: https://www.article19.org/resources/eu-code-for-practice-on-general-purpose-ai-cannot-abandon-fundamental-rights/. Per un commento, critico, della terza bozza, che andrebbe oltre il recinto dell'ambito applicativo delle disposizione del Regolamento sino ad introdurre nuove norme ancora più stringenti, si rinvia a M. Ebers, *When Guidance Becomes Overreach How the forthcoming Code of Practice Threatens to Undermine the EU's AI Act*, in *VerfBlog*, consultabile al seguente Link: https://verfassungsblog.de/when-guidance-becomes-overreach-gpai-codeofpractice-aiact/. Sulle difficoltà che sta incontrando l'Unione Europea nella implementazione dell'*AI Act* rispetto alle tempistiche originarie, si vedano R. Kilian, L. Jäck, D. Ebel, *European AI Standards - Technical Standardization and Implementation Challenges under the EU AI Act*, 2025, consultabile al seguente Link: http://dx.doi.org/10.2139/ssrn.5155591.

[34] Il testo del trattato può essere letto al seguente Link: https://rm.coe.int/1680afae3c. A commento del testo, si vedano, tra gli altri, i contributi di A. Hars, *Conceptual Difficulties in the Transformation of Human Rights to the Realm of Artificial Intelligence*, in *ACTA HUMANA*, 2024, 123 e ss.; F.P. Levantino, F. Paolucci, *Advancing The Protection Of Fundamental Rights Through AI Regulation: How The Eu And The Council Of Europe Are Shaping The Future*, in *European Yearbook on Human Rights*, 2024, Brill, in corso di pubblicazione e consultabile al seguente Link: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4881656. Per un commento sulle versioni che hanno preceduto quella poi approvata in via definitiva, si consenta il rinvio a C. Nardocci, *La (seconda) svolta del 2024. Anche il Consiglio d'Europa decide di regolamentare l'intelligenza artificiale*, in *Biolaw Journal – Rivista di biodiritto*, 2024, 73 e ss. Ancora, Sull'*iter* che ha condotto alla adozione della Framework Convention, si veda E.H. Morawska, *Council of Europe standards and activities related to AI: towards a Framework Convention on AI and human rights?*, in M. Balcerzak, J. Kapelańska-Pręgowska (a cura di), *Artificial Intelligence and International Human Rights Law*, Edward Elgar Publishing, 2024, 25 e ss.

[35] *Ad hoc Committee on Artificial Intelligence.*

[36] *Committee on Artificial Intelligence.*

fattibilità sui contenuti del trattato[37], il secondo, di predisporre le bozze di trattato[38], la Convenzione Quadro costituisce il primo trattato di diritto internazionale dei diritti umani in tema di IA avente respiro globale.

Superando la dimensione europeo-continentale, il Consiglio d'Europa ha, infatti, ammesso sin dal loro inizio la partecipazione alle negoziazioni da parte di Stati non parte. Così è stato per gli Stati Uniti, il Giappone, il Canada, Israele[39].

L'aspirazione globale della Convenzione Quadro non poggia soltanto sull'allargamento della base partecipativa alle negoziazioni. Il testo ammette, infatti, la firma e la ratifica anche da parte di Stati non parte del Consiglio d'Europa, come testimonia la firma statunitense del settembre 2024. Ma, ancora, di fronte all'inerzia oppure alla preferenza delle Nazioni Unite per strumenti di *soft law*, il Consiglio d'Europa si è imposto sulla scena quale prima ed unica organizzazione di diritto internazionale dei diritti umani ad aver messo a punto il nuovo fenomeno dell'intelligenza artificiale, disponendone il rispetto dei diritti umani fondamentali.

Venendo ai contenuti, la Convenzione Quadro presenta alcuni tratti che la accomunano all'*AI Act*, dimostrazione di una negoziazione temporalmente coeva e di reciproche influenze, ed altri che, viceversa, la allontanano, enfatizzandone la dimensione *human rights-based*.

Incidentalmente, e peraltro, vale premettere, che la distanza tra i due atti – la Convenzione Quadro, da un lato, e l'*AI Act*, dall'altro – era assai più evidente nelle prime bozze del trattato, dalla *Zero Draft* al *Consolidated Working Draft*, passando per la *Revised Zero Draft*. Differenze che si sono, poi, progressivamente assottigliate, determinando anche forse uno snaturamento del trattato del Consiglio d'Europa, nell'intento di avvicinarne il testo quanto più possibile mosso all'impianto regolatoria prescelto dal legislatore dell'Unione Europea.

Non stupisce, così, che la definizione di intelligenza artificiale, accolta dal trattato, ricalchi perfettamente quella dell'*AI Act*, ripiegando ancora una volta sulla proposta dell'OECD.

Non si muovono critiche su questo punto. Occorre, tuttavia, ricordare, che l'*endorsement* per la definizione suggerita dall'OECD rimane una scelta europea, non

---

[37] Il riferimento è al documento intitolato *Possible elements of a legal framework on artificial intelligence, based on the Council of Europe's standards on human rights, democracy and the rule of law*, pubblicato il 3 dicembre 2021 e consultabile al seguente Link: https://rm.coe.int/cahai-2021-09rev-elements/1680a6d90d. Sull'operato del CAHAI, si rinvia a A. Mantelero, *Regulating AI, in Beyond Data. Information Technology and Law Series*, T.M.C. Asser Press, The Hague, in particolare 161 e ss.

[38] Il riferimento è alla *Zero Draft* (Link: https://www.statewatch.org/media/3697/coe-artificial-intelligence-convention-zero-draft-30-6-22.pdf.), alla *Revised Zero Draft* (Link: https://rm.coe.int/cai-2023-01-revised-zero-draft-framework-convention-public/1680aa193f.), sino al *Consolidated Working Draft*, il cui testo, che risale al luglio 2023, può essere consultato al seguente Link: https://rm.coe.int/cai-2023-18-consolidated-working-draft-framework-convention/1680abde66.

[39] Nel momento in cui si scrive, la Convenzione Quadro è stata sottoscritta dai seguenti Stati: Andorra, Georgia, Islanda, Liechtenstein, Montenegro, Norvegia, Repubblica di Moldova, San Marino, Svizzera, Regno Unito, Unione Europea, Canada, Israele, Giappone, Stati Uniti.

invece universalmente condivisa, come testimoniano, gli Stati Uniti, ma anche altre realtà ordinamentali, che parlano di "sistema automatizzato", lasciando indietro la controversa nozione di intelligenza artificiale e optando per soluzioni più ampie e rispettose del principio della neutralità tecnologica.

Nonostante sia adesso condivisa la nozione di intelligenza artificiale, la Convenzione non mutua più dall'*AI* Act, come risultava dall'ultima bozza di trattato, la proposta classificatoria delle tecnologie di IA in base al criterio del rischio, viceversa assente nel testo del trattato.

Accanto alla definizione, su cui non ci si sofferma, la Convenzione ha un primo pregio: il dettaglio dei diritti nuovi o, se si preferisce, *AI-oriented*, cioè associati al funzionamento delle tecnologie di intelligenza artificiale. Non è questa la sede per soffermarsi sulla qualificazione corretta di tali diritti quali diritti nuovi oppure corollari di diritti fondamentali tradizionali. Resta che la Convenzione Quadro "fa ordine", chiarisce di quali diritti si tratta, li elenca, dà conto delle rispettive implicazioni, va a fondo nel chiarimento dei rispettivi ambiti di applicazione. Questo vale per alcuni tra quelli ripresi nel Capo III del trattato, rubricato *Principles related to activities within the lifecycle of artificial intelligence systems*, come nel caso dei principi di trasparenza e *human oversight*[40], di responsabilità e *accountability*[41], di affidabilità (*reliability*)[42], sicurezza (*safe innovation*)[43].

Al tempo stesso, la Convenzione Quadro si preoccupa di legare la nuova dimensione dell'IA ai diritti classici, con qualche rivisitazione. Interessante, per tutti, la rilettura come *personal autonomy* del diritto di autodeterminazione, quasi a voler enfatizzare l'esigenza che la persona umana non solo rimanga al centro, rievocando il nostro art. 2 Cost. e più in generale la concezione individualistico-liberale dei diritti fondamentali, ma, soprattutto, che la macchina non prenda il sopravvento sulla prima, comprimendone volontà ed azioni.

Secondo principio classico su cui si sofferma la Convenzione Quadro è, poi, la non discriminazione, evocata in due disposizioni del testo (artt. 10 e 17)[44], e valorizzata

---

[40] Cfr. Articolo 7.

[41] Cfr. Articolo 9.

[42] Cfr. Articolo 12.

[43] Cfr. Articolo 13.

[44] In letteratura, sulle relazioni tra IA e discriminazioni, si vedano F.Z. Burgesius, *Discrimination, artificial intelligence, and algorithmic decision-making*, Council of Europe Publications, 2018; J. Kleinberg, J. Ludwig, S. Mullainathan, C.R. Sunstein, *Discrimination in the Age of Algorithms*, in *Journal of Legal Analysis*, 2018, 113 e ss., e, dello stesso A.J. Kleinberg, J. Ludwig, S. Mullainathan, A. Rambachan, *Algorithmic fairness*, in *AEA papers and proceedings*, 2018, 22 e ss.; J. Gerards, *Algorithmic discrimination in Europe: Challenges and Opportunities for EU equality law*, Link: https://www.europeanfutures.ed.ac.uk/algorithmic-discrimination-in-europe-challenges-and-opportunities-for-eu-equality-law/.; P. Zuddas, *Intelligenza artificiale e discriminazioni*, in *Giurcost.*, 2020, 1 e ss.; si consenta, infine, il rinvio a C. Nardocci, *Intelligenza artificiale e discriminazioni*, in *Rivista "Gruppo di Pisa*, 2021, Link: https://www.gruppodipisa.it/images/rivista/pdf/Costanza_Nardocci_-_Intelligenza_artificiale_e_discriminazioni.pdf

anche nella sua dimensione collettiva, come testimonia il riferimento alla Convenzione Quadro per la protezione delle minoranze nazionali, qui posta al centro superando il criticabile oscurantismo, di cui si è detto, del Regolamento europeo.

Il trattato non riferisce di una discriminazione diversa, altra, da quella umana, ma, perlomeno, valorizza la questione discriminatoria facendo dell'eguaglianza e della non discriminazione uno dei principi cardine ai quali deve informarsi il funzionamento dei sistemi di intelligenza artificiale. Non solo, quindi, dichiarazioni di principio, ma vere e proprie norme a presidio dell'eguaglianza e della non-discriminazione.

Infine, una ulteriore aspetto da sottolineare riguarda il sistema di mitigazione dei rischi derivanti dal ricorso alle tecnologie di intelligenza artificiale.

La Convenzione Quadro si preoccupa delle ricadute pregiudizievoli dei sistemi di intelligenza artificiale tramite la previsione di un apposito meccanismo di valutazione di impatto che l'art. 16 della Convenzione Quadro delinea nei suoi principi guida e che il Consiglio d'Europa ha di recente assoggettato alle indicazioni concrete e dettagliate di un proprio sistema autonomo di mitigazione, HUDEIRA, predisposto dal CAI e reso noto il 28 novembre 2024[45].

Al di là delle specificità della metodologia HUDEIRA, la valutazione di impatto presupposta dall'art. 16 poggia su alcune caratteristiche, che la distanziano dall'omologa norma dell'*AI Act*. Innanzitutto, la norma chiarisce l'arco temporale di riferimento della valutazione di impatto, che, nella Convenzione Quadro, si muove a cavallo tra indagini *ex ante* ed *ex post*, a differenza dell'art. 27 *AI Act*, che pone l'accento sulla fase preventiva. In secondo luogo, e più in generale, diversa è la *ratio* che soggiace alla norma. Non una valutazione di impatto ideata in modo astratto e di generale applicazione, ma, piuttosto, un sistema che tenga conto: del contesto di impiego della tecnologia di intelligenza artificiale, delle specificità e della fase di utilizzo del sistema di intelligenza artificiale. Una valutazione di impatto, che risponde, quindi, a logiche individualizzate e caso per caso, rifuggendo applicazioni rigide, presuntivamente riferibili a qualsiasi sistema di intelligenza artificiale e che prescindono dal settore di utilizzo specifico. A differenza, infatti dal Regolamento europeo, la Convenzione non circoscrive l'ambito di operatività della valutazione *ex* art. 16 alle sole tecnologie ad alto rischio, aprendola a tutte e facendo ricadere l'obbligazione che ne deriva in capo a tutti gli Stati parte[46].

In conclusione, sebbene la Convenzione si sia progressivamente avvicinata all'impostazione e alle scelte accolte dal legislatore dell'Unione – come testimoniano le

[45] La nuova metodologia di valutazione dei rischi associati all'impiego delle tecnologie di intelligenza artificiale è pubblicata sul sito del Consiglio d'Europa e consultabile al seguente Link: https://rm.coe.int/cai-2024-16rev2-methodology-for-the-risk-and-impact-assessment-of-arti/1680b2a09f.

[46] Sulle differenze tra la Convenzione Quadro del Consiglio d'Europa e il di poco precedente AI Act, si consenta, ancora, il rinvio a C. Nardocci, *La (seconda) svolta del 2024. Anche il Consiglio d'Europa decide di regolamentare l'intelligenza artificiale*, cit. e, in particolare, della stessa A., alla monografia *Algoritmi, eguaglianza, discriminazione. Le sfide dell'intelligenza artificiale*, Giappichelli, Torino, in corso di pubblicazione.

critiche di eccessiva adesione all'*AI Act* mosse nei confronti dell'ultima versione del trattato da parte di alcune associazioni non governative[47] –, restano alcune differenze, che mantengono intatta almeno l'aspirazione originaria del Consiglio d'Europa. L'adozione, in altri termini, di un testo destinato ad avere ampia ed omogenea applicazione territoriale, stante il divieto di apporvi riserva, ed in grado di dettare i tempi e le regole del diritto internazionale dei diritti umani, con anticipo rispetto alle Nazioni Unite, nel nuovo scenario della *governance* delle tecnologie di intelligenza artificiale.

4. Un secondo: il "no" degli Stati Uniti e la nuova era della contro-regolamentazione

Se il continente europeo ha sposato sin da subito un'impostazione favorevole alla regolamentazione dei sistemi di intelligenza artificiale, l'esperienza statunitense ha conosciuto fasi alterne. Dalla reticenza originaria a normare le tecnologie di intelligenza artificiale che ha ceduto il passo ad una successiva stagione di attivismo, degli Stati, più decisa, ma anche, più morbida, dell'amministrazione federale, si è, poi, giunti sul finire del gennaio 2025 alla netta presa di posizione della nuova amministrazione che, come ripreso in apertura, si è espressa in modo chiaro contro ogni forma di regolamentazione che ostacoli la leadership statunitense nel settore dello sviluppo dei sistemi di intelligenza artificiale.

Emblematico il primo *Executive Order*, siglato dalla nuova amministrazione Trump, dal titolo *Removing Barriers to American Leadership in Artificial Intelligence*[48], che si apre appuntando l'attenzione sulla urgenza di "solidify"[49] il posizionamento degli Stati

---

[47] Con le parole "*No to the abdication of our rights*", si apriva la lettera indirizzata dalle associazioni non governative firmatarie, nel gennaio del 2024, alle istituzioni dell'Unione Europea, nonché agli Stati membri parte attiva della negoziazione affinché ostacolassero alcune delle criticità che il testo ha poi ratificato. Particolarmente significativo il passaggio conclusivo dello Statement con cui si sottolineava quanto segue: "[w]e have never given mandate to our elected representatives to abdicate our rights through the conclusion of a Convention that is ironically supposed to safeguard them. A hollowed-out Convention will provide little meaningful protections to individuals who are increasingly subject to powerful AI systems prone to bias, human manipulation, and the destabilisation of democratic institutions. As for public activities, some states are pushing for a blanket exemption with regard to national security and defence. Nothing justifies the unconditional waiving of the safeguards set in international, European and national law that usually apply in these fields. These attempts also turn a blind eye to the geopolitical context characterised by an increasing prevalence of dual use AI that put our lives and freedom at risk". il testo definitivo del trattato è stato fatto oggetto di critiche da parte di un gruppo di associazioni non governative, che hanno reso pubbliche le proprie perplessità in una *open letter* pubblicata il 5 marzo 2024. Il testo integrale può essere consultato al seguente Link: https://docs.google.com/document/d/19pwQg0r7g5Dm6_OlRvTAgBPGXaufZrNW/edit.

[48] Il testo integrale dell'Ordine esecutivo è consultabile al seguente Link: https://www.whitehouse.gov/presidential-actions/2025/01/removing-barriers-to-american-leadership-in-artificial-intelligence/.

[49] Così si esprime, letteralmente, il recente Ordine esecutivo.

Uniti quali leader globali nel settore dell'intelligenza artificiale. Urgenza, che l'*Executive Order* traduce nella revoca delle *policies* e delle direttive esistenti, ritenute un ostacolo ai danni dell'"American AI Innovation" la cui revoca risponderebbe allo scopo di "ripulire" la strada verso la *leadership* globale degli Stati Uniti in materia di intelligenza artificiale[50].

La prima azione della nuova amministrazione si muove, così, in una direzione che non è solo alternativa ma che è più propriamente opposta alla precedente, le cui azioni devono essere rimosse, "cancellate". L'ordine esecutivo non contiene riferimenti espressi a quali siano tutte le *policies* e le direttive oggetto della revoca.

Accanto all'esplicito riferimento all'ordine esecutivo disposto dall'amministrazione precedente nel 2023, *Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence*, un chiarimento sarà, probabilmente, fornito dall'*action plan* che un insieme di soggetti elencati nel nuovo ordine esecutivo dovrà presentare al Presidente degli Stati Uniti entro i prossimi 180 giorni.

L'interrogativo aperto riguarda, in definitiva, quale sarà l'impatto che il nuovo ordine esecutivo ed l'*action plan* di futura definizione saranno in grado di produrre in relazione alla proposta di legge, attualmente all'esame del Congresso, e che costituisce il punto più avanzato raggiunto dalla *Biden-Harris administration* espressione, al tempo stesso, di un approccio chiaramente ispirato a logiche di regolamentazione. Logiche di regolamentazione, a cui nettamente si oppone la nuova amministrazione la cui linea di azione appare ben più rappresentata dall'espressione "contro-regolamentazione", ad enfatizzare il superamento della *ratio* della *self-regulation* e lo spostamento verso il rifiuto di qualsiasi norma che si proponga di irregimentare il funzionamento delle tecnologie di intelligenza artificiale.

4.1.Prima del "no": l'approccio olistico dell'amministrazione federale precedente

Se il 2025 ha inaugurato la stagione della "contro-regolamentazione", gli anni precedenti avevano testimoniato una impostazione diversa e disponibile ad inserirsi nel seminato tracciato, per prima, dall'Unione Europea.

Sino alla fine del 2024, l'esperienza statunitense presentava, così, svariati profili di interesse, che merita ricordare, in questa sede, in primo luogo, perché espressiva di una modalità alternativa a quella europea di assicurare la *governance* delle tecnologie di intelligenza artificiale. In seconda istanza, perché, se è indubbia, come si è detto poco sopra, la *ratio* del provvedimento della nuova amministrazione, non del tutto chiari ne

[50] Un secondo e più recente Ordine esecutivo è stato siglato il 23 aprile 2025, dal titolo *Advancing Artificial Intelligence Education for American Youth*, il cui testo integrale è consultabile al seguente Link: https://www.whitehouse.gov/presidential-actions/2025/04/advancing-artificial-intelligence-education-for-american-youth/. L'Ordine esecutivo si propone di promuovere e rafforzare la *AI literacy and proficiency* nel settore dell'istruzione attraverso l'istituzione di una *Task Force on Artificial Intelligence Education*, che dovrà favorire l'impiego delle nuove tecnologie di intelligenza artificiale quale strumento educativo ad ampia diffusione all'intero degli istituti scolastici.

sono gli effetti nei confronti di tutti gli atti adottati dalla precedente e su cui ci si sofferma, seppure brevemente, qui di seguito[51].

Procedendo a ritroso, ci si riferisce alla prima proposta di legge, il *Federal Artificial Intelligence Risk Management Act*, sottoposta all'attenzione del Congresso nel 2024, all'*Executive Order* dell'amministrazione *Biden-Harris* dell'autunno 2023, sino al *Blueprint for an AI Bill of Rights* dell'ottobre del 2022[52].

Partendo dal *Blueprint*, il documento, ancorché non vincolante, presentava svariati profili innovativi.

Il primo era rappresentato dall'approccio olistico alle tecnologie di intelligenza artificiale. A differenza, cioè, dell'impostazione continentale-europea, che centra l'*AI Act* sui rapporti tra persona e sistemi di intelligenza artificiale, il *Blueprint* ne valorizzava all'opposto la dimensione collettiva con una presa di posizione, che costituiva la novità più importante nel panorama globale degli interventi regolatori.

L'approccio olistico traspariva in modo evidente dalle sezioni dedicate all'opportunità, che il funzionamento dei sistemi di intelligenza artificiale si mantenga entro binari conformi al principio di parità. Si leggeva, così, che "the harms of automated systems should be evaluated, protected against, and redressed *at both the individual and community levels* [*corsivo nostro*]"[53], e che il concetto di comunità, su cui indugia il testo, "is integral to the scope of the *Blueprint for an AI Bill of Rights*"[54]. Affermazioni, che dimostrano l'esigenza di comprendere e contrastare i fenomeni, quello discriminatorio per primo e più volte evocato nel testo – altra differenza di peso rispetto all'impostazione europea –, anche a livello collettivo. Il testo ritornava, così, più volte, sul concetto di comunità, affiancandovi l'elenco delle c.d. *underdeserved*

---

[51] Sull'esperienza statunitense come sviluppatasi sino alla conclusione del 2024 si consenta il rinvio a C. Nardocci, *Dalla "self-regulation" alla frammentata regolamentazione dei sistemi di intelligenza artificiale: uno sguardo alla (diversa) prospettiva statunitense*, in Rivista di diritto pubblico comparato ed europeo, 2024, 859 e ss.

[52] Il testo integrale del *Blueprint* può essere consultato al seguente link: https://bidenwhitehouse.archives.gov/wp-content/uploads/2022/10/Blueprint-for-an-AI-Bill-of-Rights.pdf.

Sui contenuti del *Blueprint*, si vedano, in letteratura, E. Himme, L. Floridi, *The Blueprint for an AI Bill of Rights: In Search of Enaction, at Risk of Inaction*, in *Minds and Machines*, 2023, 285 e ss.; E. Lance, *US Releases an AI Bill of Rights That, Though Encouraging, Won't Yet Move the Needle*, 2022, reperibile al seguente Link: https://www.jurist.org/commentary/2022/10/u-s-releases-an-ai-bill-of-rights-that-though-encouraging-wont-yet-move-the-needle/; M. Amarikwa, *Rules for robots: constitutional challenges with the ai bill of right's principles regulating automated systems*, in *Journal of Constitutional Law*, 2024, 1176 e ss.; L. Eliot, *US Releases an AI Bill of Rights That, Though Encouraging, Won't Yet Move the Needle*, *JURIST – Commentary — Legal News & Commentary*, consultabile al seguente Link: https://www.jurist.org/commentary/2022/10/u-s-releases-an-ai-bill-of-rights-that-though-encouraging-wont-yet-move-theneedle/, 2022.

[53] Così, il *Blueprint*, cfr. 10.

[54] *Ibidem*.

communities, cioè le vittime dei nuovi sistemi di intelligenza artificiale[55]; sull'impatto collettivo di tali sistemi; sull'esigenza di operare valutazioni comparative, in termini di sotto oppure di sovra-rappresentanza, ancora una volta, su basi collettive; su una concezione di discriminazione, che la avvicina a quella statistica, di cui discorrono le scienze economiche e che pone a raffronto gruppi più che individui; da ultimo, su alcuni diritti, come il diritto di opzione (*to opt out*), ricollegato al diritto "nuovo" alla alternativa umana, azionabile dinanzi a decisioni della macchina sprovviste del previo consenso della persona.

Accanto all'impostazione di principio e, con essa, alla valorizzazione del dato collettivo, possono essere ricordate almeno le scelte, che lo separavano dalla prospettiva regolatoria europea.

La prima è la definizione di intelligenza artificiale, a cui il *Blueprint* preferiva a nozione di "sistema automatico". La definizione presenta un impianto a maglie larghe, che non distingueva tra tecnologie più o meno avanzate, sistemi *machine o deep learning* e semplici software, e che poneva, piuttosto, l'accento sulla capacità lesiva di tutti questi strumenti, "automatici", sui diritti della persona.

Secondo punto di disallineamento (il secondo) rispetto all'*AI Act*, era costituito dall'omesso riferimento al rischio quale criterio per classificare le tecnologie di intelligenza artificiale.

La categorizzazione delle tecnologie di IA in base alla rispettiva – ma predeterminata in astratto – capacità lesiva non trovava accoglimento entro l'ordinamento statunitense. Per il *Blueprint*, ciò che conta, ai fini del ricorso legittimo ad una tecnologia di intelligenza artificiale, è la sua non nocività sul piano del suo impatto sui diritti fondamentali.

Anche in relazione alla valutazione di impatto, che, come detto, costituisce il perno dell'intervento regolatorio europeo, si scorge qualche frizione. Astenendosi dalla individuazione di strategie specifiche, il *Blueprint* optava per un'impostazione *bottom up* e demandava a meccanismi partecipativi dal basso, cioè a consultazioni pubbliche, l'identificazione dei meccanismi più efficaci di mitigazione dei rischi. Altrettanto interessante è il suggerimento per la istituzione di meccanismi indipendenti di

[55] Il *Blueprint* dedica una sezione specifica alla nozione di "underserved community", cfr. 11. Della nozione di comunità, inoltre, il documento offre una definizione propria, secondo cui "[c]ommunities" include: neighborhoods; social network connections (both online and offline); families (construed broadly); people connected by affinity, identity, or shared traits; and formal organizational ties. This includes Tribes, Clans, Bands, Rancherias, Villages, and other Indigenous communities. AI and other data-driven automated systems most directly collect data on, make inferences about, and may cause harm to individuals. But the overall magnitude of their impacts may be most readily visible at the level of communities. Accordingly, the concept of community is integral to the scope of the Blueprint for an AI Bill of Rights. United States law and policy have long employed approaches for protecting the rights of individuals, but existing frameworks have sometimes struggled to provide protections when effects manifest most clearly at a community level".

controllo, aprendo la partecipazione a tale processo ad un gruppo eterogeneo di soggetti – ricercatori, giornalisti, revisori terzi –, ai quali assicurare un accesso tempestivo ai dati strumentali all'espletamento delle funzioni di supervisione. Spunto, che potrebbe essere coerente con l'art. 77 dell'*AI Act*, se non fosse che la norma europea non pare incline ad aprire l'inclusione nella "lista" delle autorità di vigilanza a così un così vasto gruppo di soggetti.

Vi era, infine, un quarto aspetto innovativo. Ci si riferisce, così, all'adesione esplicita alla tesi secondo cui la discriminazione derivante dai sistemi di intelligenza artificiale costituirebbe una fenomenologia discriminatoria eterogenea rispetto alla c.d. *human-driven discrimination*, discorrendo di *algborithmic discrimination*. In questo senso, i passaggi dedicati al *proxy* e ai *proxies*, che, più che i fattori di discriminazione tradizionali, sono alla base delle decisioni dei sistemi di intelligenza artificiale, e che il *Blueprint* richiedeva fossero assoggettati ad un sindacato in tutto analogo a quello che la legge riserva ai *suspect grounds of discrimination*[56].

Il *Blueprint* non ha costituito, però, il solo prodotto dell'attivismo federale della passata amministrazione, ad oggi oggetto del già esaminato provvedimento di revoca.

Nell'ottobre del 2023, l'amministrazione *Biden-Harris* aveva, infatti, pubblicato un proprio *Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence*[57], che demandava al Dipartimento di Stato il compito di redigere la prima guida per il settore pubblico e privato, la società civile, le organizzazioni non governative, tesa alla definizione di alcuni criteri per valutare l'impatto dei sistemi di IA sui diritti fondamentali, poi resa pubblica nel luglio del 2024. Si tratta del *Risk Management Profile for Artificial Intelligence and Human Rights*, altrimenti denominato "*The Profile*".

Sul piano degli avanzamenti e deviazioni dal *Blueprint* e, soprattutto, dal Regolamento europeo, meritano di essere ricordate: la messa a punto di un programma, che sviluppi tecnologie avanzate di intelligenza artificiale in gradodi garantire la sicurezza *cyber*; il contrasto della discriminazione algoritmica, sviluppando i portati di un precedente ordine esecutivo del febbraio 2023, *Executive Order on Further Advancing Racial Equity and Support for Underserved Communities Through The Federal Government*;

---

[56] Non a caso il documento intitola un paragrafo "Guarding against proxies", cfr. 26.

[57] Il testo integrale dell'*Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence* può essere letto al seguente Link: https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/. A commento del testo, si rinvia a M. Wörsdörfer, *Biden's Executive Order on AI: strengths, weaknesses, and possible reform steps*, in *AI Ethics*, 2024; e, dello stesso A., e per un confronto con l'*AI Act* dell'Unione Europea, si veda, *Biden's Executive Order on AI and the E.U.'s AI Act: A Comparative Computer-Ethical Analysis*, in *Philos. Technol.*, 2024, 73 e ss.; D.D. Douglas, *Responsibilities of an Executive Leading AI Projects: Navigating Federal Directives for Safe and Inclusive Development*, Preprints, 2024, il testo può essere consultato al seguente Link: https://www.preprints.org/manuscript/202402.0815/v1.

l'individuazione di azioni concrete a supporto della regolamentazione responsabile ed effettiva dei sistemi di intelligenza artificiale[58].

Da ultimo, si colloca il *Federal Artificial Intelligence Risk Management Act*, che si proponeva di imporre l'utilizzo dell'*Artificial Intelligence Risk Management Framework*, sviluppato dal *National Institute of Standards and Technology* (NIST)[59], ogni qualvolta le agenzie federali decidessero di fare uso di sistemi di intelligenza artificiale.

L'incorporazione del modello di valutazione del rischio entro una proposta di legge federale costituiva l'approdo indubbiamente più significativo, soprattutto dalla prospettiva europea dove ancora si attende di comprendere: quali saranno le linee direttrici concrete della valutazione di impatto presupposta dall'art. 27 dell'*AI Act*; chi – oltre l'attività in corso in seno alla Commissione UE – sarà incaricato di redigere le prime, allo scopo di assicurarne una omogeneità di utilizzo.

Questi, in estrema sintesi, gli approdi della precedente amministrazione, in evidente contrasto con la "contro-regolamentazione" del nuovo corso.

4.2.E le scelte di campo statali: da e oltre il Colorado

Ben più dirompente e dinamico anche del primo approccio federale e pure a fronte delle nuove tendenze imposte a livello federale, continua ad essere lo scenario statale.

Se difettano discipline che si preoccupino di affrontare in modo organico tutti i profili, sono almeno 17 gli Stati, che hanno ad oggi adottato discipline preposte alla regolamentazione di qualche aspetto del ricorso alle tecnologie di intelligenza artificiale.

In estrema sintesi, sono tre gli approcci normativi che si scorgono a livello statale. In prima istanza, si inseriscono gli Stati (quelli intervenuti per primi), che hanno optato per una disciplina dedicata in via esclusiva alle interazioni tra intelligenza

[58] Più in particolare, due azioni venivano indirizzare alle aziende, rese destinatarie, da un lato, di interventi di *guidance*, che includessero standard di condotta chiari, protettivi dei diritti fondamentali e delle esigenze di *safety* in sede di programmazione e sviluppo dei sistemi di intelligenza artificiale; dall'altro, di un supporto nell'acquisizione di beni e servizi di intelligenza artificiale, che sia rapido, economico ed efficiente. Una terza azione, infine, riservata ad una vasta cerchia di *stakeholders*, riguardava, invece, la formazione di professionisti in materia di intelligenza artificiale, di cui veniva incaricato l'*Office of Personnel Management*.

[59] Il modello elaborato dal NIST, che avrebbe dovuto costituire il corrispettivo del prossimo *Code of Practice* in corso di adozione dalle istituzioni dell'Unione Europea, articolava, e tuttora articola, la valutazione del rischio in quattro fasi: la mappatura dei sistemi di intelligenza artificiale, che le organizzazioni intendono impiegare tenendo conto dei vantaggi offerti rispetto ad altre metodiche, anche umane; la misurazione, che attesti i rischi associati al sistema di intelligenza artificiale prescelto; la fase di gestione dei rischi identificati, tramite un monitoraggio costante, che copra rischi prevedibili, ma anche imprevisti; la delineazione di *policies* e processi che garantiscano l'utilizzo safe and secure del sistemi di intelligenza artificiale.

artificiale, diritto alla riservatezza e protezione dei dati (*Privacy-Based Approach*), come nel caso di: California, Colorado, Connecticut, Delaware, Indiana, Iowa, Montana, Oregon, Tennessee, Texas e Virginia. Una seconda impostazione è, invece, propria, degli Stati, che hanno preferito regolamentazioni settoriali (*Sectoral Approach*), volte a normare l'impiego dei sistemi di intelligenza artificiale entro ambiti specifici, come quello lavoristico (Illinois)[60] oppure dei *deepfake* (Tennessee)[61]. Da ultimo, venendo al terzo approccio (*Human Rights-Based Approach*), si collocano gli Stati che si sono preoccupati di disciplinare le relazioni tra intelligenza artificiale e principi fondamentali, con attenzione alle ripercussioni della prima sull'eguaglianza e la non discriminazione.

Prima di passare al Colorado, protagonista indiscusso di quest'ultimo approccio regolatorio, non mancano anche esempi di Stati sulla scena statunitense, che hanno introdotto norme in tema *accountability*. Regolamentazioni, cioè, che hanno scelto di inserire programmatore e utilizzatore (e non solo il secondo) nel novero di coloro che potrebbero essere chiamati a rispondere in presenza di un eventuale malfunzionamento del sistema di intelligenza artificiale. Tra queste, si segnala, in particolare, il Tennessee, che, in tema di profilazione, ha introdotto l'obbligo di condurre valutazioni di impatto in capo a chi sviluppa, chi implementa e, infine, a chi fa uso di sistemi di intelligenza artificiale, rifiutando l'impostazione, prevalente sul continente europeo, che concentra in capo all'ultimo soggetto della catena ogni profilo di *liability*.

Ma, come premesso, è il Colorado ad avere conquistato la scena, approvando nel 2024 la prima regolamentazione integralmente dedicata al contrasto della *algboritmic discrimination*, *Colorado Artificial Intelligence Act* (CAIA)[62].

Primo esempio di approccio capillare al fenomeno della discriminazione *AI-based*, il CAIA affronta i due temi più cruciali e controversi del dibattito attuale, legando a doppio filo violazione dei diritti, intelligenza artificiale e responsabilità (*liability*), individuale oppure congiunta, del programmatore, dell'utilizzare, e di qualsiasi parte terza che si interfacci con il primo oppure con il secondo. Quanto alla dimensione discriminatoria, il CAIA non affronta questioni di diritto anti-discriminatorio stretto,

[60] Il riferimento è alla proposta HB 3773, che si occupa in modo specifico del contrasto alla discriminazione algoritmica sul luogo di lavoro, approvata nel giugno del 2024.

[61] Sul Tennessee, si segnala, in particolare, la legge approvata nel 2023 dallo Stato del Tennessee, che ha introdotto norme specifiche in tema di profilazione, prescrivendo l'obbligo di condurre valutazioni di impatto tanto in capo a chi sviluppa, quanto a chi implementa e fa uso di sistemi di IA. Sempre lo Stato del Tennessee ha, poi, nel marzo del 2024, approvato una seconda normativa, l'*Ensuring Likeness, Voice and Image Security (ELVIS) Act*. La legge, primo esempio di disciplina preposta ad offrire una forma di salvaguardia in particolare dei musicisti a livello statunitense, è entrata in vigore il primo luglio 2024. Il testo integrabile può essere letto al seguente link: https://legiscan.com/TN/text/HB2091/id/2900923.

[62] Il testo può essere consultato al seguente Link: https://leg.colorado.gov/sites/default/files/2024a_205_signed.pdf.

omettendo di esprimersi in ordine alla piena cittadinanza degli istituti e delle categorie classiche anche alla discriminazione *AI-based*.

Tuttavia, senza ricorrere alle nozioni di *proxy* e di *proxy discrimination*[63], il testo fornisce la prima proposta definitoria di discriminazione *AI-based* a livello globale.

La discriminazione algoritmica è, cioè, qualificata alla stregua di qualsiasi condizione o modalità tramite cui l'utilizzo di una tecnologia di intelligenza artificiale si traduce in un trattamento oppure in un impatto irragionevolmente differenziato che pregiudica un individuo oppure un gruppo di individui – e, qui, vi è un passaggio decisivo – sulla base di caratteristiche attuali o presunte. Nozione che ha il pregio di incardinare la definizione, seppure implicitamente, sulla bipartizione discriminazione diretta / indiretta; e, forse, soprattutto, di non ripiegare su un elenco non esaustivo di categorie protette, cioè qualità individuali oppure collettive, effettive e presenti (*actual*), preferendo affiancarvi il riferimento a caratteristiche presunte (*perceived)*, cioè associate al singolo (e al gruppo) oppure, ancora, ritenute presuntivamente sussistenti e che si siano rivelate, però, decisive ai fini del trattamento oppure dell'effetto discriminatorio prodotto dal sistema di intelligenza artificiale.

Il secondo aspetto che regge l'impianto del CAIA attiene, invece, ai profili di responsabilità, per la prima volta, interpretati in senso ampio sia dal punto di vista soggettivo, dei soggetti potenzialmente *liable,* sia da quello oggettivo, cioè dei contenuti delle obbligazioni di programmatore e utilizzatore.

Non è questa la sede per passare in rassegna tutte le obbligazioni che il CAIA pone, partitamente, in capo ai due soggetti. Va, però, segnalata la scelta di rendere potenzialmente responsabili sia il programmatore che l'utilizzatore finale, in costanza di una discriminazione *AI-based*, richiedendo ad entrambi di procedere ad una valutazione di impatto del sistema di intelligenza artificiale. La condivisione

[63] Per un approfondimento sulla nozione di "proxy discrimination", si rinvia a A.E.R. Prince, D. Schwarcz, *Proxy discrimination in the age of artificial intelligence and big data*, in *Iowa Law Journal*, 1277 e ss.; A. Datta et al., *Proxy Discrimination in Data-Driven Systems: Theory and Experiments with Machine Learnt Programs*, 2017, in https://arxiv. org/pdf/1707.08120.pdf.; J. Grimmelmann, D. Westreich, *Incomprehensible Discrimination*, in *California Law Review*, 2017, 164 e ss.; L. Alexander, K. Cole, *Discrimination by Proxy*, in *Constitutional Commentary*, 1997, 453 e ss. Si consenta, inoltre, il rinvio a C. Nardocci, *Proxy Discrimination in Artificial Intelligence: What We Know and What We Should Be Concerned About*, in https://www.chairesante.ca/en/articles/2024/proxy-discrimination-in-artificial-intelligence-what-we-know-and-what-we-should-be-concerned-about/, 2024. Più in generale, sulle specificità della "AI-based discrimination", si vedano F.Z. Burgesius, *Discrimination, artificial intelligence, and algorithmic decision-making*, Strasburgo, 2018; S. Barocas, A.D. Selbst, *Big data disparate impact*, in *California Law Review*, 2016, 671 e ss.; J. Kleinberg, J. Ludwig, S. Mullainathan, C.R. Sunstein, *Discrimination in the Age of Algorithms*, in *Journal of Legal Analysis*, 2018, 113 e ss., dello stesso A., anche, J. Kleinberg, J. Ludwig, S. Mullainathan, A. Rambachan, *Algorithmic fairness*, in *AEA papers and proceedings*, 2018, 22 e ss.; si consenta, infine, il rinvio a C. Nardocci, *Intelligenza artificiale e discriminazioni*, in *Rivista "Gruppo di Pisa*, 2021 e, della stessa A., si rinvia alla monografia *Algoritmi, eguaglianza, discriminazione. Le sfide dell'intelligenza artificiale*, Giappichelli, Torino, in corso di pubblicazione.

dell'obbligazione di effettuare la valutazione di impatto tra programmatore e utilizzatore, costituisce tratto caratterizzante la disciplina, che la distanzia dal modello europeo, che, viceversa, predilige la responsabilità esclusiva dell'utilizzatore demandando a quest'ultimo l'obbligo la valutazione di impatto. Collegata alle obbligazioni di programmatore e utilizzatore, il CAIA costruisce, infine, anche una nuova azione giudiziaria, su cui non ci si sofferma, azionabile dopo la notifica con cui l'utilizzare comunica al Procuratore generale la scoperta degli effetti discriminatori di una tecnologia di intelligenza artificiale.

Per concludere, il CAIA costituisce la disciplina che ha saputo, per prima, prendere sul serio la discriminazione algoritmica, dandovi un nome e attribuendo responsabilità ad una estesa gamma di soggetti, che interagiscono con la macchina. Il CAIA si affranca, così, dal silenzio dell'*Executive Order* federale, dal timido approccio europeo e, anche, dalle prime risposte giurisprudenziali, più inclini a considerare responsabile l'utilizzatore, facendo retrocedere corresponsabilità del programmatore.

All'opposto, tale ampliamento delle responsabilità anche al programmatore va accolto con particolare favore, perché risponde ai tempi e alle caratteristiche della discriminazione *AI-based*, le cui origini e manifestazioni sono complesse ed intricate, così come le relazioni tra tutti coloro che, a vario titolo, contribuiscono a fare entrare nello spazio umano le tecnologie di IA e favorirne, talvolta, gli annessi effetti pregiudizievoli.

Da ultimo, vale ricordare che il Colorado non è rimasto isolato[64].

Altri Stati, tra cui l'Illinois[65], il Connecticut[66], la Virginia[67], la California[68] e, più di recente, il Texas[69] e il New Jersey[70] ne hanno seguito le orme, sia sotto il versante dell'esplicito contrasto alla discriminazione *AI-based*, sia quanto all'attenzione prestata alla *liability* tramite il dettaglio delle obbligazioni e delle responsabilità di tuti i soggetti

---

[64] Valga segnalare, da ultimo, che una legge successiva (H 1468) ha istituito una *Task Force AI*, suggerendo modifiche al CAIA prima della sua entrata in vigore effettiva, prevista nel 2026.

[65] Il riferimento è alla proposta HB 3773, che si occupa in modo specifico del contrasto alla discriminazione algoritmica sul luogo di lavoro, approvata nel giugno del 2024. A questa si affianca la proposta SB2203, rubrica *Preventing Algorithmic Discrimination Act*, presentata nel febbraio 2025.

[66] Il testo della proposta SB2, presentata il 2 aprile 2024 ed approvata nel gennaio 2025, rubricata *An Act Concerning Artificial Intelligence*, è reperibile al seguente Link: https://www.cga.ct.gov/2024/ba/pdf/2024SB-00002-R01-BA.pdf.

[67] Così il Bill H.R. n. 2094 del 2025.

[68] Si veda, in questo senso, la proposta AB n. 1018 del 2025.

[69] Il riferimento è al Bill HB n. 1709, presentato il 23 dicembre 2024.

[70] Il riferimento è alla recente approvazione e pubblicazione della *Guidance on Alghorithmic Discrimination and the New Jersey Law Against Discrimination*, del gennaio 2025, consultabile al seguente Link: https://www.nj.gov/oag/newsreleases25/2025-0108_DCR-Guidance-on-Algorithmic-Discrimination.pdf, con cui lo Stato del New Jersey ha esteso la normativa statale in tema di diritto anti-discriminatorio anche nei confronti della discriminazione generata dalle tecnologie di intelligenza artificiale.

coinvolti: dal programmatore sino all'utilizzatore finale. A livello municipale, si segnala, infine, il caso pioneristico della città di New York[71].

5.Un caso a sé: una "contro-regolamentazione" alternativa: il Canada e l'AI Countergovernance

Dopo aver preso per anni sul serio le implicazioni giuridiche dell'intelligenza artificiale, un altro ordinamento, che ha fatto un passo indietro e che fa parlare di "contro-regolamentazione", o meglio, di *AI-countergovernance*[72], è il Canada.

La storia del Canada presenta qualche affinità con le relazioni tra livelli di governo, che ha caratterizzato la prima esperienza statunitense. Anche in Canada, si è assistito, così, ad un primigenio movimento dal basso, che non ha interessato le province – fatta salva l'eccezione del Québec[73] –, ma piuttosto alcune realtà locali. Il riferimento è alle città di Montréal e di Toronto, protagoniste della stesura di documenti pioneristici – la *Montréal Declaration on Responsible AI*[74] e la *Toronto Declaration*[75] –, fondative del c.d. approccio *human-centered AI*, esportato sulla scena globale.

Più controversa e deludente è stata, invece, l'esperienza a livello federale, dove l'*Artificial Intelligence and Data Act* (AIDA), prima proposta di legislazione in tema di intelligenza artificiale, in discussione presso la *House of Commons* dal giugno 2022 e vera l'alternativa all'*AI Act*, è stato stralciato ad inizio 2025 insieme alle dimissioni rilasciate dal Primo Ministro Justin Trudeau.

La vicenda, che ha visto protagonista l'AIDA, merita di essere segnalata, poiché espressiva di una tendenza, non nuova[76], tradotta nel concetto teorico, ripreso in

---

[71] Il riferimento è alla proposta di legge *Local Law* 144, entrata in vigore nel luglio 2023. L'ambito applicativo della normativa è circoscritto alle discriminazioni causate da tecnologie di intelligenza artificiale, che si verifichino sul luogo di lavoro.

[72] Il termine è stato, per la prima volta, coniato da J.D. Rikki, nel suo *Counter-Governance: Citizen Participation Beyond Collaboration*, in *Cogitatio*, 2018, 180 e ss.

[73] Per una panoramica approfondita delle iniziative intraprese dal Québec in tema di intelligenza artificiale e, in particolare, di *governance*, si rinvia al sito ufficiale *Responsible AI in Québec* di cui al seguente Link: https://ia.quebec/en/governance.

[74] La *Montréal Declaration on Responsible AI* è stata adottata e resa pubblicata nel 2018. Il testo può essere letto al seguente link: https://declarationmontreal-iaresponsable.com/wp-content/uploads/2023/04/UdeM_Decl-IA-Resp_LA-Declaration-ENG_WEB_09-07-19.pdf

[75] La *Toronto Declaration. Protecting the right to equality and non-discrimination in machine learning systems* è stata adottata nel maggio 2018 ed è consultabile al seguente link: https://www.torontodeclaration.org/declaration-text/english/.

[76] Un caso interessante di *AI countergovernance* si è verificato nel 2018 e ha avuto Google come protagonista. La vicenda nasceva dalle proteste di un gruppo di lavoratori, che si era opposto alla collaborazione avviata da Google con il Pentagono per lo sviluppo di un sistema di IA da impiegare nel settore militare per il funzionamento di droni. L'opposizione dei lavoratori ha costretto Google ad abbandonare il progetto (*Project Maven*). Per questa notizia, si veda il seguente Link:

apertura, di *AI-countergovernance*. Con l'espressione *AI-countergovernance* si vorrebbe, cioè, definire quel processo tramite il quale le comunità, i lavoratori e le lavoratrici, si oppongono alle iniziative di *governance* dell'intelligenza artificiale promosse dai governi e dal mondo dell'industria, poiché ritenute lesive dei loro diritti e non corrispondenti ai loro interessi[77].

L'*Artificial Intelligence and Data Act* (AIDA) si allontanava dall'impostazione europea per alcuni aspetti cruciali, al centro di quelle stesse critiche che ne hanno, poi, determinato il fallimento. Per prima, la definizione di intelligenza artificiale: l'AIDA si occupava piuttosto soltanto dei c.d. *high impact systems*, omettendo di fornire un definizione, e circoscrivendo il proprio ambito applicativo in modo particolarmente stringente. Una seconda differenza rispetto all'*AI Act* riguardava, poi, il riferimento al rischio, quale criterio classificatorio astratto, sostituito nella proposta canadese dal grado di impatto (*high impact*) – inteso come probabilità che una tecnologia di intelligenza possa incidere negativamente sui diritti fondamentali –, senza, però, che venisse chiarito il "come" dovesse essere effettuata simile valutazione.

Una terza deviazione rispetto all'approccio riguardava la scelta del governo canadese di assoggettare alla regolamentazione in esame soltanto il settore privato e non anche il pubblico. E, così, non a caso, qualche convergenza con il continente europeo si scorgeva in relazione alle tecnologie di intelligenza artificiale fatte oggetto di divieto e, in particolare, alle deroghe al loro impiego relative ad esigenza di sicurezza

---

https://gizmodo.com/google-plans-not-to-renew-its-contract-for-project-mave-1826488620. Ancora, nel 2019, nella città di Toronto, si è verificato un altro esempio di opposizione ad un progetto avente ad oggetto lo sviluppo di una tecnologia di intelligenza artificiale che, in questo caso, avrebbe dovuto essere impiegata in un quartiere della città (*Sidewalk Toronto project*), denominata *Smart city*, da parte di una impresa denominata *Sidewalk Labs*, una azienda "sorella" di Google. Una comunità di residenti del quartiere *(#BlockSidewalk*) aveva espresso le proprie preoccupazioni sulla gestione dei dati e sul sistema di controllo e di sorveglianza pubblica presupposti dal progetto. La resistenza della comunità cittadina si è tradotta in alcuni incontri pubblici e nella presentazione di progetti alternativi sino alla cancellazione del progetto da parte della *big tech* coinvolta nel 2020, all'indomani della diffusione della pandemia da Covid-19. Nonostante la concomitanza con l'emergenza sanitaria, alcuni osservatori (Cfr. i commenti consultabili a questo Link: https://onezero.medium.com/how-a-band-of-activists-and-one-tech-billionaire-beat-alphabets-smart-city-de19afb5d69e) hanno sottolineato l'influenza decisiva del movimento di *backlash* o di opposizione promosso dalla comunità locale rispetto alla decisione di ritirare il progetto.

[77] Si tratta di un fenomeno che si contrappone alla intelligenza artificiale partecipativa o *Participatory AI*, che, viceversa, presuppone il coinvolgimento diretto e l'*empowerment* di coloro che risultino direttamente esposti a subire le conseguenze del ricorso ai sistemi di intelligenza artificiale sulla base dell'assunto che tale coinvolgimento favorisca un funzionamento più equo, inclusivo e *trustworthy* delle tecnologie di intelligenza artificiale. Per un approfondimento in letteratura, si vedano A. Birhane, W. Isaac, V. Prabhakaran, M. Díaz, M.C. Elish, I. Gabriel, S. Mohamed, *Power to the People? Opportunities and Challenges for Participatory AI*, in *Equity and Access in Algorithms, Mechanisms, and Optimization (EAAMO '22)*, 2022, consultabile al seguente Link: https://dl.acm.org/doi/10.1145/3551624.3555290.

e di sorveglianza pubblica[78]. Ci si riferisce ai sistemi di identificazione biometrica, utilizzati senza il previo consenso della persona per ragioni di identificazione oppure di valutazione di attitudini comportamentali; le tecnologie utilizzate allo scopo di stabilire chi abbia o meno diritto di accedere ad un servizio; i sistemi impiegati in ambito sanitario e dell'amministrazione della giustizia.

Su questi tre aspetti, in particolare, si sono sviluppate le prese di posizione contrarie all'approvazione dell'AIDA[79]. Tuttavia, le ragioni dell'opposizione nei confronti della proposta di legge hanno investito soprattutto il metodo, che ha guidato il governo federale nel corso dell'*iter* parlamentale, che rende la *AI-countergovernance* molto diversa da quella che si è, in questa sede, denominato come "contro-regolamentazione", cioè il fenomeno che sta caratterizzando l'approccio statunitense.

Se, negli Stati Uniti, la *ratio* degli interventi recenti – in sintesi la revoca dell'ordine esecutivo della amministrazione *Biden-Harris* – mira a soddisfare le esigenze delle maggiori imprese tecnologiche, che auspicano per un settore poco regolamentato e libero da norme restrittive, la *AI-countergovernance* canadese si ispira a ragioni diametralmente opposte. Sarebbe stata, infatti, l'esclusione dalla discussione parlamentare di una vasta gamma di soggetti e comunità – quelle esposte al rischio di subire gli effetti pregiudizievoli di un impiego delle tecnologie di intelligenza artificiale regolamentato in modo lasco – a fondare il movimento contrario all'approvazione dell'AIDA.

A fronte della massiccia consultazione di *stakeholders* provenienti dal settore *tech*[80], la stessa attenzione e disponibilità di coinvolgimento, viceversa, non sarebbe stata riservata ad altre categorie, tra cui, in particolare, le popolazioni indigene, gruppi etnici minoritari, organizzazioni di *advocacy* e non governative sindacati, le realtà locali.

La *AI-countergovernance*, che ha conosciuto in Canada una traduzione concreta, denuncia, quindi, l'omesso allargamento della base partecipativa e, quale conseguenza, la ridotta sensibilità prestata dal tentativo regolatorio nei confronti della tutela dei diritti fondamentali, avendo viceversa optato per una concentrazione del dibattito politico a senso unico tra le istituzioni governative federali, da un lato, l'industria *tech*, dall'altro.

---

[78] Critiche a questo specifico aspetto dell'AIDA sono oggetto, tra gli altri, alla *cautionary note* redatta dal *Canadian Civil Liberties Association*, consultabile al seguente Link: https://ccla.org/wp-content/uploads/2023/09/Bill-C-27-Submission-to-INDU-CCLA.pdf.

[79] In tema, si veda, anzitutto, il Report redatto da C. Tessono, Y. Stevens, M.M. Ma-Lik, S. Solomun, S. Dwivedi, S. Andrey, *AI Oversight, Accountability and Protec-ting Human Rights. Cybersecure Policy Exchange*, 2022, consultabile al seguente link*: https://www.cybersecurepolicy.ca/aida*. Una sintesi efficace delle principali criticità dell'AIDA è offerta dal Montréal AI Ethics Institute, consultabile al seguente link: *https://montrealethics.ai/the-death-of-canadas-artificial-intelligence-and-data-act-what-happened-and-whats-next-for-ai-regulation-in-canada/*.

[80] Su cui si veda A. Clement, *Preliminary analysis of ISED's C-27 list of 300 stakeholder consultation meetings*, 2003, consultabile al seguente Link: https://ssrn.com/abstract=4658004, che presenta una analisi critica delle audizioni intraprese dal governo federale nei primi due anni di *iter* parlamentare dell'AIDA.

6.Un terzo. La via di mezzo di chi, dopo l'Unione Europea, vorrebbe regolamentare: il Brasile, la Corea del Sud, la Cina

Un terzo approdo dei tempi più recenti, e che si sta sempre più consolidando, riguarda le numerose proposte di regolamentazione delle tecnologie di intelligenza artificiale che, da più parti, sono state presentate e che si trovano all'esame dei rispettivi organi legislativi.

Non potendo analizzarle tutte, si è scelto di considerare, in questa sede, alcune tra quelle che si trovano in uno stato più avanzato del rispettivo *iter* di approvazione e che presentano tratti innovativi se poste a raffronto con l'*AI Act* europeo.

Tra queste, si colloca il Brasile dove, nel dicembre 2024, il Senato ha approvato la proposta di legge n. 2338/2023, che costituirà la prima normativa in tema di intelligenza artificiale del Paese. La proposta sarà sottoposta, nel corso del 2025, all'esame della Camera dei Deputati, prima della firma del Presidente che ne segnerà la definitiva entrata in vigore.

La proposta brasiliana contiene alcuni punti di contatto con l'impostazione europea.

Anche il Bill n. 2338, come l'*AI Act*, adotta un approccio fondato sul rischio, distinguendo tra tecnologie di intelligenza artificiale vietate (*excessive risk)*, poiché ree di causare pregiudizi eccessivi ai diritti fondamentali, "ad alto rischio" (*high risk*), da assoggettare ad una serie di requisiti di utilizzo aggiuntivi e altre suscettibili di determinare rischi di altra natura (*other risk*). A differenza, però, del Regolamento europeo, la proposta di legge brasiliana non limita la valutazione di impatto alle tecnologie di intelligenza artificiale ad alto rischio, estendendone l'ambito di applicazione nei confronti di tutti i sistemi. Una scelta da guardare con favore, poiché attenua le criticità che caratterizzano classificazioni astratte e rigide, consentendo una disamina delle ricadute pregiudizievoli anche di tecnologie, originariamente non considerate rischiose, ma suscettibili di diventarlo per effetto di applicazioni future e non conoscibili a priori.

Ancora, differisce dalle scelte europee, almeno nella misura in cui omette di ripiegare sulla nozione suggerita dall'OECD, la definizione di intelligenza artificiale. Secondo la proposta di legge brasiliana, per sistema di intelligenza artificiale si deve, infatti, intendere: "a computational system, with varying degrees of autonomy, designed to infer how to achieve a given set of objectives, using approaches based on machine learning and/or logic and knowledge representation, through input data from machines or humans, with the aim of producing predictions, recommendations, or decisions that may influence the virtual or real environment"[81]. Una nozione, che appare più ampia e non limitata alle tecnologie più sofisticate, come si verifica nel Regolamento europeo.

[81] Così recita l'art. 4, par. n. 1.

Un altro punto di divergenza riguarda i principi che soggiacciono alla *governance* delle tecnologie di intelligenza artificiale. A questo riguardo, ancorché implicitamente ripresi dall'*AI Act*, la proposta di legge brasiliana presenta un grado di dettaglio maggiore, tale da rendere azionabili e giustiziabili i principi in essa sanciti. Tra questi, poi, ve ne sono almeno due che, se oscurati nel Regolamento europeo, conoscono, invece, nella contesto della proposta brasiliana una maggiore e meritata evidenza: l'*accountability* o *liability*, da un lato; e la non discriminazione, dall'altra, resa principio cardine delle valutazioni di impatto.

Sul versante della implementazione e dell'*enforcement*, la proposta di legge demanda ad un organismo regolatorio *ad hoc*, affiancato da un *Executive Branch*, il rispetto delle sue prescrizioni attraverso meccanismi di *auditing*, *reporting* ed attività investigative aventi ad oggetto denunciate violazioni delle sue norme.

In conclusione, la proposta costituisce un esempio meritorio di *governance*, che coniuga le *best practices* emerse a livello globale con nuove strategie di gestione delle nuove tecnologie di intelligenza artificiale che riflettono le specificità locali.

Se il Brasile rappresenta un caso interessante per i contenuti di quella che rimane, ad oggi, ancora una proposta, un posto di primo piano va riservato alla Corea del Sud, che dopo l'Unione Europea e la Cina, ha approvato definitivamente una propria regolamentazione in materia di intelligenza artificiale a fine gennaio 2025.

Si tratta della c.d. "*AI Basic Law*", legislazione che presenta alcune affinità, ma anche differenze importanti se posta a raffronto con il Regolamento europeo.

Come già l'*AI Act*, la *AI Basic Law* adotta un approccio in base al rischio, prevedendo una serie di obbligazioni per i *deployers* dei sistemi di intelligenza artificiale definiti – con una differenza terminologica – "ad alto impatto". Ricorre, quindi, anche nella normativa sudcoreana una preferenza per assoggettare alle procedure di *risk assesment* soltanto le tecnologie di intelligenza artificiale ritenute potenzialmente più lesive.

Ancora, simile all'*AI Act* è la scelta delle deroghe all'impiego di determinate tecnologie di intelligenza artificiale, laddove, ancora una volta, l'utilizzo è considerato legittimo laddove risponda ad esigenze di sicurezza e di sorveglianza pubblica.

Maggiori sono, però, le differenze tra le due normative che muovono, anzitutto, dalla eterogeneità di approccio.

La *AI Basic Law* presenta una impostazione meno incline a tracciare limiti e divieti e più favorevole a sostenere l'innovazione, seguendo un modello che si sta affermando sempre di più nel corso degli ultimi mesi. In questa stessa direzione, come si vedrà, si sta muovendo, ad esempio, il Regno Unito, in linea con quello che ormai si definisce il "*wait and see approach*"; un approccio alla regolamentazione delle nuove tecnologie di intelligenza artificiale, che ritiene i tempi non ancora maturi i tempi per normative stringenti, preferendovi soluzioni flessibili e adattabile agli sviluppi ed implementazioni future delle prime.

Sempre sul versante degli elementi di distinzione, e allontanandosi dalla linea regolatoria europea, la *AI Basic Law* non contiene alcuna proposta classificatoria delle tecnologie di intelligenza artificiale – ivi incluse di quelle vietate, di cui non vi è alcuna tipizzazione –, optando, all'opposto, per un approccio di tipo settoriale. In altri termini, la *AI Basic Law* definisce tecnologie di intelligenza artificiale "ad alto impatto" quei sistemi suscettibili di causare un danno ad uno o più diritti fondamentali se impiegati in alcune delle aree o ambiti di utilizzo, questi sì tipizzati per legge. Tra i settori che rendono "ad alto impatto" una tecnologia di intelligenza artificiale, si collocano, allora: l'ambito lavoristico e assicurativo quando i sistemi di intelligenza artificiale siano impiegati per determinare assunzioni oppure concessioni di credito; il settore dell'istruzione laddove faccia ricorso a sistemi di IA allo scopo di fornire valutazioni di studenti e studentesse; l'amministrazione della giustizia, qualora faccia ricorso a sistemi di identificazione facciale per disporre un arresto oppure per condurre delle indagini.

Una ulteriore differenza attiene, poi, alle modalità di realizzazione della valutazione di impatto, o meglio, ai soggetti coinvolti nella medesima. La *AI Basic Law* non la demanda a parti terze – come, viceversa, prevede l'*AI Act* –, ma stabilisce che la valutazione venga disposta sotto l'egida del Ministero della scienza e delle tecnologie dell'informazione e della comunicazione. Valutazione di impatto che costituisce, altro elemento di eterogeneità, un obbligo non solo per il *deployer*, bensì anche per il programmatore o sviluppatore della tecnologia di intelligenza artificiale.

Infine, merita segnalare che, a differenza dell'AI Act che contiene una serie di obblighi di compliance per le sole tecnologie "ad alto rischio", la *AI Basic Law* vi assoggetta tutti i sistemi di intelligenza artificiale. Una scelta che costituisce il risvolto immediato del rifiuto di classificare a monte le tecnologie di intelligenza artificiale e che meglio risponde al principio della neutralità tecnologica di quanto non accada per la impostazione rigidamente top-down del Regolamento europeo.

Da ultimo, un cenno merita la Cina, unico ordinamento giuridico che ha sostenuto sin da subito insieme all'Unione Europea l'opportunità di regolamentare i nuovi sistemi di intelligenza artificiale[82].

A differenza dell'Unione Europea che ha optato per una regolamentazione organica ed unitaria, la Cina ha, viceversa, inizialmente, demandato la *governance* dei sistemi di intelligenza artificiale a testi normativi differenziati per oggetto ed ambito applicativo. Il prodotto dell'approccio settoriale (*sector-specific*) ha, così portato all'adozione di tre normative, qui riportate nella loro versione inglese: *Regulations on the Management of Algorithm Recommendations for Internet Information Services*, *Regulations on Deep*

[82] Approfondisce le opzioni regolatorie della Repubblica Popolare Cinese rispetto alle soluzioni vagliate dall'Unione Europea, Generare e non creare? Spunti per una comparazione sulla regolazione dell'intelligenza artificiale generativa tra Stati Uniti, Repubblica Popolare Cinese e Unione Europea D. Clementi, *Generare e non creare? Spunti per una comparazione sulla regolazione dell'intelligenza artificiale generativa tra Stati Uniti,*

Repubblica Popolare Cinese e Unione Europea, in *Rivista di Diritti Comparati*, 2024, 371 e ss.

Integration Management of Internet Information Services e, da ultimo, le *Interim Measures for the Management of Generative Artificial Intelligence Services*.

Nel maggio 2024, invece, è stata resa nota la prima bozza della prima proposta di legge in materia di intelligenza artificiale, la c.d. *Artificial Intelligence Law of the People's Republic of China*[83].

In estrema sintesi, questi i punti salienti della proposta.

In primo luogo, l'approccio che dovrà ispirare la prossima e prima legislazione cinese in materia di intelligenza artificiale è "*people's centered*", invece che ispirato a logiche individualistiche e chiara espressione dell'impianto socialista della Repubblica cinese. Le finalità della nuova legge non si limitano, però, a porre il "popolo" al centro, assicurando la tutela dei diritti fondamentali. La proposta cinese esplicitamente sancisce al suo primo articolo, che obiettivo della prossima regolamentazione sarà coniugare innovazione tecnologica e sviluppo dei nuovi sistemi di intelligenza artificiale con la imprescindibile salvaguardia dei diritti fondamentali.

A questo riguardo, merita segnalare che la proposta di legge contiene un elenco dettagliato e assai più esaustivo, di quanto non accada invece per il Regolamento europeo, dei principi cardine in materia di intelligenza artificiale. Una scelta, quella di rendere espliciti i diritti "nuovi" e meno nuovi, che supera l'opzione accolta dalle istituzioni dell'Unione Europea non soltanto per la previsione di un Capo *ad hoc*, il Capo Terzo, a questi dedicato, ma anche per la riconosciuta rilevanza assegnata a diritti viceversa non esplicitati oppure trascurati dall'*AI Act*. Si muovono in questo senso: il diritto all'intervento umano, ma, ancora di più, quello alla contestazione di decisioni assunte in via esclusiva dalla macchina (art. 35); la centralità riconosciuta al principio di eguaglianza e di non discriminazione (parte prima; art. 32); la espressa previsione di una norma interamente dedicata ai diritti e agli interessi di coloro che versino in condizioni di particolare svantaggio (art. 38), come i minori, le persone in età avanzata, le donne, le persone con disabilità; la previsione di una norma specifica a tutela dei diritti e degli interessi dei lavoratori.

Per quanto attiene, invece, alle scelte compiute in merito alle (eventuali) tecnologie da vietare oppure da rendere destinatarie di controlli più stringenti, la proposta di legge cinese non introduce alcun elenco di sistemi di intelligenza artificiale vietati, né una classificazione tesa a tracciare i confini tra queste ultime, quelle qualificate "ad alto rischio" e le restanti.

Questo approccio, che si pone coerentemente in linea con le più recenti proposte e legislazioni in materia, non esclude obblighi di verifica della sicurezza dei sistemi di intelligenza artificiale, che la proposta di legge cinese configura in capo ai *providers* e ai *deployers* estendendone, poi, l'ambito applicativo nei confronti di tutte le tecnologie senza alcuna distinzione. Ancora una volta, quindi, un'opzione regolatoria che si distanzia dal modello europeo anche sul piano delle linee direttrici a cui si informa la

[83] La traduzione in inglese della prima bozza può essere consultata al seguente Link: https://cset.georgetown.edu/wp-content/uploads/t0592_china_ai_law_draft_EN.pdf.

valutazione del rischio, inserendosi, piuttosto, nel solco di quelle maggioritarie nel dibattito attuale, alcune delle quali hanno, come visto, conosciuto inveramento concreto, come nel caso della esaminata *AI Basic Law* sudcoreana.

6.1.Il Regno Unito, che parte dai "principi"

Il posizionamento del Regno Unito nella costruzione dicotomica regolamentazione / non regolamentazione delle tecnologie di intelligenza artificiale – sia essa "contro-regolamentazione" oppure *AI-countergovernance* – non è agevole.

Anche per il Regno Unito, il 2025 ha segnato una svolta.

Sebbene non avesse mai sposato una impostazione integralmente favorevole alla regolamentazione dei sistemi di intelligenza artificiale sulla scia dell'*AI Act*, nel corso degli ultimi anni, il Regno Unito aveva, tuttavia, posto in essere una serie di iniziative che sembravano preludere alla prossima adozione di una legislazione organica avente ad oggetto la *governance* delle tecnologie di intelligenza artificiale.

In questa prospettiva, un primo documento da segnalare è costituito dal *White Paper*, pubblicato nel 2023, con cui il Governo conservatore ha tratteggiato in modo molto chiaro la propria *ratio* ispiratrice: coniugare innovazione tecnologica, da un lato, e contenere i rischi derivanti dalle tecnologie di intelligenza artificiale, rispondendo alle preoccupazioni della società, dall'altro.

Il *White Paper*, palesava, inoltre, la presa di posizione del Regno Unito contraria all'adozione in tempi brevi di una regolamentazione, fatta di norme rigide e stringenti, sulla scorta del Regolamento dell'Unione Europea, considerata *rea* di porre vincoli eccessivi alle imprese, limitando la libera espressione dello sviluppo tecnologico e frustrando l'innovazione. All'opposto, il *White Paper* proponeva di prendere le mosse dalla un insieme di principi agili, seguendo un approccio *pro-innovation*[84], per valutare in una fase successiva l'opportunità, eventuale, di demandare al Parlamento la definizione di una legislazione in materia.

La flessibilità costituisce, quindi, elemento cardine del documento, a cui si accostano l'enucleazione di principi guida – secondo una impostazione che non è solo *pro-innovation*, ma, anche, *principles-based approach* – ed un sistema di supervisione e monitoraggio in tempo reale delle modalità di utilizzo delle tecnologie di intelligenza artificiale da parte delle imprese, in grado di intercettare rischi emergenti e imprevisti.

[84] In senso conforme alla impostazione che si sta delineando nel Regno Unito, si stanno muovendo anche altri ordinamenti giuridici, che preferiscono un approccio morbido alle tecnologie di intelligenza artificiale, che non si traduca in norme stringenti, come, viceversa, vengono interpretate le norme dell'*AI Act* approvato dalle istituzioni dell'Unione Europea. In questo senso, si stanno indirizzando il Giappone, Singapore, l'Australia, la Svizzera. Avevano pronosticato che l'"effetto Bruxelles" avrebbe condotto alla sperimentazione di vie alternative alla regolamentazione delle nuove tecnologie di intelligenza artificiale M. Almada, A. Radu, *The Brussels Side-Effect: How the AI Act Can Reduce the Global Reach of EU Policy*, in *German Law Journal*, 2024, 646 e ss.

Nelle intenzioni degli estensori del *White Paper*, tanto i primi e, in particolare, il secondo avrebbero dovuto consentire un adattamento morbido e graduale del quadro normativo, allo scopo di tenere in adeguata considerazioni le specificità di ciascuna tecnologie di intelligenza artificiale e le implicazioni del loro utilizzo in dipendenza del settore considerato.

Il *White Paper* si caratterizzava, poi, per l'adesione ad un modello di tipo collaborativo, fondato, cioè, sul coinvolgimento diretto delle imprese nella definizione dei suoi contenuti ed implementazione concreta, secondo una logica che mirava a valorizzare il dialogo, trilaterale, tra Governo, "*regulators*" e imprese.

L'impostazione del *White Paper* non è mai stata sinora abbandonata e, anzi, le più recenti iniziative del Governo laburista si allineano più con questo approccio, che non con la spinta regolatoria che sembrava aver preso il sopravvento nel corso della prima metà del 2024. Facevano, infatti, parlare di una prossima "legge" del Regno Unito in tema di intelligenza artificiale, anzitutto, la pubblicazione del documento intitolato *Initial Guidance for Regulators*, che sviluppava gli intendimenti del *White Paper* e, in particolare, la prima fase di guida dei *regulators* funzionale all'approvazione di una regolamentazione vera e propria; l'istituzione, nel 2024, dell'*AI Office* ed il suo inquadramento presso il Dipartimento per la Scienza, l'Innovazione e la Tecnologia, attivo sin dal 2021 e incaricato di redigere e aggiornare le *AI National Strategies*; ma, soprattutto, la stesura di una proposta di legge vera e propria, la prima, inizialmente sottoposta alla *House of Lords* dal precedente governo conservatore: la *Artificial Intelligence (Regulation) Private Members' Bill (the "Bill")*.

Il 2025 si è, viceversa, aperto con due eventi che destano qualche perplessità rispetto agli intendimenti reali del Regno Unito verso l'adozione di una legislazione in materia di intelligenza artificiale: l'omessa firma, unico Stato insieme agli Stati Uniti, dello *Statement* che ha concluso i lavori del *Summit di Parigi*, da un lato; la re-introduzione, prima, della proposta di legge in tema di intelligenza artificiale (*The Bill*) presso la *House of Commons*, ed il posticipo (forse) della sua adozione definitiva, dall'altro. Forse, poiché notizie di stampa riportano, che vi sarebbe l'intenzione del Governo laburista di ritardare l'*iter* parlamentare della proposta di legge, approvata in prima lettura presso la *House of Lords* nel mese di marzo 2205, per favorire l'allineamento del Regno Unito con la nuova impostazione statunitense.

Quale sarà l'esito dell'esame parlamentare è difficile prevederlo.

Certamente, il Regno Unito ha proposto sino ad oggi un modello di gestione delle tecnologie di intelligenza artificiale alternativo. Alternativo alla "contro-regolamentazione" statunitense, poiché espressivo, all'opposto, della volontà di delineare un quadro normativo, ancorché con tempi e modalità graduali e frutto di un modello dialogico-partecipativo (riservato ad imprese e *regulators*); ma anche alla regolamentazione "pura" dell'*AI Act*, di cui rifiuta l'astrattezza in punto di classificazione delle tecnologie di intelligenza artificiale e di qualificazione dei rischi ed annessi divieti.

Tutto questo basta per giustificare il difficile posizionamento, di cui si dava conto in apertura. Resta, così, da attendere l'esito del dibattito parlamentare, ma, anche, verificare i contenuti della (eventuale) legge che diranno quanto il Regno Unito sia vicino o distante, anzitutto, dall'Unione Europea.

7.Un quarto. La dimensione sovranazionale: le altre organizzazioni internazionali regionali, le Nazioni Unite e la preferenza per strumenti di soft law

Il Consiglio d'Europa non è la sola organizzazione internazionale ad occuparsi dei rapporti tra intelligenza artificiale e diritti fondamentali. Le più importanti organizzazioni internazionali sia globali che regionali sono, infatti, quasi tutte intervenute in materia e con tempi e modalità differenti.

Tra queste, vanno ricordate: l'UNESCO, a lungo impegnata sui temi dell'etica dell'intelligenza artificiale e della promozione di una visione inclusiva, soprattutto di genere e non discriminatoria dei sistemi di intelligenza artificiale[85]; e una serie di organizzazioni internazionali regionali, tra cui l'organizzazione degli Stati americani (OAS), la Commissione africana per i diritti dell'uomo e dei popoli e l'organizzazione degli Stati del Sudest asiatico (ASEAN).

Venendo per prime alle iniziative promosse dalle tre principali organizzazioni internazionali e guardando agli approdi più recenti, quanto, per prima, all'organizzazione degli Stati americani (OAS), merita segnalare la dichiarazione, *Towards The Safe, Secure, and Trustworthy Development and Deployment of Artificial Intelligence in the Americas: the Importance of Governance, Regulatory, and Institutional Frameworks*[86], e l'*Action Plan*[87], adottati all'esito dell'incontro, svoltosi il 13 e il 14 dicembre 2024, tra i Ministri ed altre autorità competenti in materia di scienza e tecnologia degli Stati membri dell'OAS.

---

[85] Il riferimento, in particolare, è alla prima *Recommendation on the Ethics of Artificial Intelligence* a livello globale, adottata nel 2021 e consultabile al seguente Link: https://unesdoc.unesco.org/ark:/48223/pf0000381137. Note e meritorio è, inoltre, l'impegno profuso dall'UNESCO nel settore dell'istruzione tramite l'approfondimento delle sue relazioni con le tecnologie intelligenza artificiale, così come l'attenzione posta sulle implicazioni di genere del ricorso all'intelligenza artificiale. Su quest'ultimo punto, si ricorda il noto report *I'd blush if I could: closing gender divides in digital skills through education*, consultabile al seguente Link: https://unesdoc.unesco.org/ark:/48223/pf0000367416. Sulle relazioni tra intelligenza artificiale e genere si rinvia a M. D'Amico, C. Nardocci, *Discriminazioni, Donne e Intelligenza Artificiale*, in *AI Anthology*, Il Mulino, Bologna, 2022.

[86] Il testo della dichiarazione può essere consultato al seguente link: https://www.oas.org/en/media_center/press_release.asp?sCodigo=E-100/24.

[87] Al seguente link il testo: http://scm.oas.org/IDMS/Redirectpage.aspx?class=XVIII.7CIDI/REMCYT-VII/doc&classNum=5&lang=e.

Aspetto saliente del documento è costituito dall'invito rivolto agli Stati di promuovere forme di regolamentazione delle tecnologie di intelligenza artificiale a livello domestico, che tengano conto delle specificità del singolo contesto nazionale e dei settori di impiego delle prime. Accanto al *favor* per la regolamentazione, si affianca un secondo passaggio interessante che attiene al metodo. Sulla scorta della c.d. *Participatory AI*, il documento dell'OAS insiste affinché sia promosso il dialogo tra le istituzioni degli Stati, il mondo dell'industria *tech*, le organizzazioni non governative e la società civile, al fine di assicurare uno sviluppo delle tecnologie di intelligenza artificiale che si dimostri il quanto più possibile pertinente e coerente rispetto alle istanze che provengono dai settori direttamente coinvolti nei processi di sviluppo dei nuovi sistemi oppure che di queste subiscono o sono esposti al rischio di subirne conseguenze lesive dei loro diritti.

Non si scorgono, invece, indicazioni sui contenuti, né sui temi su cui si snoda il dibattito domestico e sovranazionale attuale: dalla definizione di intelligenza artificiale, all'eventuale classificazione delle relative tecnologie, passando dalle modalità di accertamento dei rischi, per giungere poi ai divieti da opporre all'utilizzo di taluni di tali sistemi.

In definitiva, una dichiarazione che si arresta ai principi, e che incrocia e risponde a due questioni preliminari del dibattito sul *se* e sul *come* regolamentare le nuove tecnologie di intelligenza artificiale: da un lato, l'OAS manifesta una chiara preferenza per forme di *governance* nazionali, escludendo l'adozione di un trattato proprio e speculare alla Convenzione Quadro del Consiglio d'Europa; dall'altro, l'organizzazione supporta un approccio alla regolamentazione che si fonda sullo schema dell'intelligenza artificiale partecipativa (c.d. *Participatory AI*), mettendosi, così, al riparo dalle critiche che hanno contraddistinto la descritta esperienza canadese, il cui fallimento è dipeso in larga parte, come detto, se non integralmente dall'omesso coinvolgimento di tutti i soggetti interessati.

Quanto al continente africano, accanto ad alcune iniziative domestiche, nel 2024, il Consiglio Esecutivo dell'Unione degli Stati africani ha adottato la *Continental AI Strategy*[88]. La *Continental AI Strategy* costituisce un documento articolato, che contiene una serie di indicazioni rivolte, partitamente, agli Stati, al settore privato ed agli sviluppatori delle tecnologie di intelligenza artificiale e che si propone di accelerare lo sviluppo delle nuove tecnologie anche entro il continente africano.

A tali fini, la strategia offre una propria definizione di intelligenza artificiale[89], alternativa a quelle europee e nord-americane, e si sofferma in modo dettagliato sui

[88] Il testo della *AI Continental Strategy* è consultabile in lingua inglese al seguente link: https://au.int/sites/default/files/documents/44004-doc-EN-_Continental_AI_Strategy_July_2024.pdf

[89] Così si legge nel testo: "[t]here is no universal definition of Artificial Intelligence. Within the framework of this Strategy, AI refers to computer systems that can simulate the processes of natural intelligence exhibited by humans where machines use technologies that enable them to learn and adapt, sense and interact, predict and recommend reason and plan, optimise procedures and parameters,

rischi associati all'impatto delle nuove tecnologie sui diritti fondamentali. Con riferimento a questi ultimi, ed accanto a quelli noti e ripresi dai coevi documenti legislativi e di *soft law*, la strategia ne menziona espressamente alcuni, viceversa, posti in secondo piano o, almeno, non così centrali sulla scena occidentale: i rischi ambientali e climatici, i rischi definiti "strutturali"; il contrasto con i valori continentali, i c.d. "valori africani" (*african values*)[90], taluni dei quali – la coesione sociale, la tutela dei diritti umani e la democrazia – di respiro globale, altri, quelli legati alla salvaguardia del patrimonio culturale e delle tradizioni e costumi indigeni, viceversa di specifica rilevanza continentale. Se non tutto è una novità, lo è, almeno, l'attenzione dedicata ai rischi inclusi nella prima categoria, cioè quelli qui definiti "globali": il *job displacement*, l'*AI divide*, la discriminazione di genere.

La *Continental AI Strategy* indugia, poi, sull'importanza della normazione o *governance* delle tecnologie di intelligenza artificiale, che si traduce in azioni da promuovere a livello domestico, tramite la modifica di leggi esistenti e l'introduzione di nuove tese a colmare i "regulatory gaps"[91], senza escludere azioni sovranazionali-continentali volte ad assicurare la coerenza degli interventi messi in atto ai vari livelli. Su questo ultimo punto e a differenza dell'OAS, la *African Union* si attribuisce una funzione di supervisione e di coordinamento allo scopo di assicurare il rispetto della strategia da parte degli Stati membri.

Infine, e non potendo in questa sede approfondirne tutti gli aspetti, la *Continental AI Strategy* segue l'impostazione europea sul piano della scelta di classificare, seppure in modo autonomo, le tecnologie di intelligenza artificiale e sulla imprescindibilità della valutazione del loro impatto sui diritti fondamentale quale strumento di supervisione e controllo.

Altrettanto interessanti e concreti sono, invece, gli sviluppi, che provengono dalla organizzazione degli Stati del Sud-est asiatico (ASEAN).

---

operate autonomously, be creative and extract knowledge from large amounts of data to make decisions and recommendations", 14.

[90] Si tratta dei valori elencati nel documento, cfr. 25-26 e tra cui si collocano i seguenti: "Societal cohesion […]. Democracy and Human Rights […]. Subversion of Indigenous Knowledge and African Cultural Heritage [precisando, in proposito, che] "AI systems risk appropriating and misrepresenting indigenous knowledge, potentially eroding cultural heritage and perpetuating cultural exploitation. Additionally, inadequate representation of indigenous practices in AI models may lead to the marginalisation of indigenous communities and their knowledge systems that have always been part and parcel of African norms and values".

[91] In tre passaggi del documento, si insiste su un approccio che vada a colmare le lacune esistenti nella regolamentazione dei sistemi di intelligenza artificiale. In particolare, il testo si riferisce a "[…] labour protections for gig and platform workers, standards for the public procurement of AI systems, and regulatory approval of AI for use as medical devices within health systems. Regulations related to social media, content generators and content distributors should be aligned with international normative frameworks and standards", 32-33.

Sul finire del 2024, l'ASEAN ha, infatti, adottato una propria guida regionale, pubblicata nel mese di dicembre, dal titolo *AI governance and ethics*[92].

La guida ambisce a diventare il primo documento atto a fornire un insieme di best practices funzionale a garantire e promuovere un approccio omogeneo ed un utilizzo ispirato a principi comuni delle tecnologie di intelligenza artificiale impiegate dagli Stati e dalle organizzazioni operanti nel Sud-est asiatico.

Si tratta di un documento di *policy*, che abbraccia una impostazione morbida al fenomeno dell'intelligenza artificiale che ricorda l'approccio del Regno Unito e che si contrappone, viceversa, alle scelte compiute dall'Unione Europea e, poi, confluite nell'*AI Act*. Numerosi sono gli aspetti di rilievo della Guida. Aspetti che distanziano il c.d. *light touch approach* dell'ASEAN dall'Unione Europea, ma anche da alcune soluzioni vagliate a livello domestico, come per gli Stati Uniti dell'era *pre* e *post* Trump.

Se ne ricordano, in questa sede, due, che interessano due profili su cui più carente, se non addirittura lacunoso, si dimostra il Regolamento europeo. Il primo è costituito dall'enfasi posta sul tema della l*iability*, posta non soltanto in capo al *deployer*, ma, secondo una logica che rifletta il ciclo di creazione e di vita delle tecnologie di intelligenza artificiale, anche il programmatore e il responsabile della messa in funzione del sistema di intelligenza artificiale. In secondo luogo, si colloca la proposta di modello di *risk assesment*, accluso alla guida e denominato "*risk assesment template*". Il modello delineato dalla guida è costruito sulla base di una lista di domande, con risposte chiuse e/o aperte, destinate agli sviluppatori e agli utilizzatori di sistemi di intelligenza artificiale, inclusi comitati etici ed organismi preposti alla *governance* dell'intelligenza artificiale entro organizzazioni.

Venendo, invece, alle Nazioni Unite, l'intervento della principale organizzazione internazionale si è fatto attendere. Risale solo al marzo del 2024 la pubblicazione della prima Risoluzione in tema di intelligenza artificiale, la *Draft Resolution on Safe, Secure and Trustworthy Artificial Intelligence System(s)*[93], seguita da altro documento di pari rilievo, il *White Paper on AI Governance*[94]. Ispirato dalla volontà di fornire alcune linee guida tese ad orientare il governo dell'intelligenza artificiale, il *White Paper* delinea in modo chiaro i rischi associati al ricorso ai sistemi di intelligenza artificiale a cui vi associa alcune raccomandazioni rivolte agli Stati.

L'individuazione e la selezione dei rischi associati all'intelligenza artificiale costituisce l'aspetto sicuramente più innovativo del testo rispetto alle scelte coeve ed europee, in particolare.

---

[92] Il testo è consultabile in lingua inglese al seguente link: https://asean.org/wp-content/uploads/2024/02/ASEAN-Guide-on-AI-Governance-and-Ethics_beautified_201223_v2.pdf.

[93] La notizia e i contenuti della Risoluzione possono essere consultati al seguente link: https://www.google.com/search?client=safari&rls=en&q=Un+Safe%2C+Secure+and+Trustworthy+Artificial+Intelligence+System(s)&ie=UTF-8&oe=UTF-8.

[94] Il testo può essere consultato al seguente link: https://unsceb.org/sites/default/files/2024-11/UNSystemWhitePaperAIGovernance.pdf.

Il *White Paper* descrive i rischi globali, attuali e futuri, delle tecnologie di intelligenza artificiale prendendo le mosse dal dato ontologico, dagli elementi, cioè, che ne caratterizzano la struttura, il funzionamento e gli effetti verso l'esterno. Il riferimento è, così, alla natura decentralizzata, orizzontale, multisettoriale dei sistemi di intelligenza artificiale. Il *White Paper* insiste, così, sulla circostanza per la quale i sistemi di IA presuppongono il coinvolgimento di una pluralità di soggetti, ciascuno deputato ad assolvere ad una funzione nelle varie fasi di sviluppo del sistema di intelligenza artificiale. La natura decentralizzata e, potrebbe dirsi, la partecipazione di più attori costituisce, ad avviso dell'ONU, la prima tra le caratteristiche che rendono rischioso l'impiego delle tecnologie di intelligenza artificiale sul versante, duplice: della corretta identificazione del/dei responsabile/i della violazione di un diritto fondamentale; e dell'accesso al giudice della vittima.

Ma le specificità dei sistemi di intelligenza artificiale non si esaurirebbero nella decentralizzazione. La multi-settorialità, intesa come compresenza di una molteplicità di discipline, rappresenta infatti la seconda caratteristica rea di acuire i rischi legati all'impiego delle tecnologie di intelligenza artificiale. Più sono i settori e le aree di interesse, maggiori saranno le conseguenze, eventualmente, pregiudizievoli generate dai sistemi di intelligenza artificiale.

Alla decentralizzazione e alla multi-settorialità dei sistemi di intelligenza artificiale, il *White Paper* associa alcuni rischi, poco discussi sulla scena europea: la proliferazione di principi nuovi, a cui non corrisponderebbe una altrettanto chiara definizione dei procedimenti di accertamento delle violazioni l'identificazione del responsabile; il ruolo degli attori non statali, che sposterebbero il baricentro decisionale al di fuori delle istituzioni pubbliche e rappresentative con i connessi rischi di privatizzazione; l'omessa inclusione di categorie e gruppi sociali svantaggiati nel dibattito sulla *governance* dell' intelligenza artificiale. Il *White Paper* specifica ulteriormente il terzo rischio appena richiamato, individuandone almeno due cause: la contrapposizione tra Nord e Sud del mondo, legata alla ineguale distribuzione delle risorse e delle *expertise* tecnologiche; le relazioni, altrettanto ineguali, inter-statali; la condizione delle minoranze etniche, delle donne e dei gruppi sociali LGBTQ+, non rappresentati ed esclusi nelle fasi di costruzione e messa in funzione dei sistemi di intelligenza artificiale.

Questa ricostruzione dei rischi, interessante e nuova perché ne emancipa la lettura da una impostazione tecnica per suggerirne una rivisitazione che ne valorizzi la dimensione giuridica e sociale, si affianca ad una serie di raccomandazione che, complessivamente intese, palesano un *favor* deciso per l'approccio *human rights-based*, sintetizzabile nella richiesta di rendere ogni regolamentazione in materia di intelligenza artificiale rispettosa del diritto internazionale dei diritti umani. Un'impostazione, poi, accolta dalla citata *Draft Resolution on Safe, Secure and Trustworthy Artificial Intelligence System(s)*.

Anche nella Risoluzione, le Nazioni Unite si occupano di temi meno arati nel dibattito europeo: l'allargamento della partecipazione alla trasformazione digitale in favore di tutti e, soprattutto dei Paesi in via di sviluppo per ricomporre il *digital divide* tra e all'interno degli Stati; la riduzione degli altri *gaps* generati dai sistemi di intelligenza artificiale sul piano delle relazioni inter-individuali.

Viene, allora, da interrogarsi sulle ragioni del silenzio europeo. Se la ridotta eterogeneità socio-economica tra gli Stati membri potrebbe motivare la posizione delle istituzioni euro-unitarie, non sembra di potersi dire lo stesso per il Consiglio d'Europa, che, viceversa, conosce e fronteggia le conseguenze delle diseguaglianze tra Ovest ed Est Europa. Peraltro, non si ritiene che il silenzio su questo tema costituisca solo il prodotto della ridotta disparità economica e sociale tra gli Stati membri delle due organizzazioni. Piuttosto, verrebbe da dire, come scarso lo spazio, poi in parte riconosciuto, ai principi di eguaglianza e di non discriminazione tra individui, così, similmente, difetta il recepimento pieno di questi stessi principi sul piano dei rapporti tra gruppi e, in ultima analisi, tra Stati.

La preoccupazione per l'omessa inclusione dei Paesi del Sud del mondo non è prerogativa delle Nazioni Unite. Analoga impostazione è stata condivisa da altre organizzazioni di diritto internazionale regionale – come l'organizzazione degli Stati americani (OAS) e, parimenti, simili spunti si scorgono nella discussione dell'*High-Level Forum on Artificial Intelligence (AI)*[95], svoltasi entro il continente africano nel luglio 2024 –, così come da alcune realtà domestiche, poco sopra approfondite, come il Brasile.

8. Riflessioni conclusive: si sta (ancora) tra assonanze e dissonanze

L'analisi che precede ha evidenziato come gli approcci alle tecnologie di intelligenza artificiale, da una convergenza recente e temporalmente collocabile sul finire del 2024, si siano nuovamente distanziati, come dimostra la c.d. "contro-regolamentazione", sposata in modo emblematico dal caso statunitense, espressione di una vera e propria insofferenza nei confronti della normazione o governance delle nuove tecnologie.

Pure a fronte di questa separazione, che caratterizza i rapporti tra gli Stati Uniti dell'amministrazione Trump e, per certi aspetti anche il Canada, da un lato, l'Europa e il resto del mondo, almeno sinora, dall'altro, pare interessante porre l'accento su alcuni punti fermi, che contraddistinguono le regolamentazioni attuali o in discussione sia in termini di convergenza che di dissonanza.

Partiamo dagli elementi di convergenza.

---

[95] Il riferimento è al primo *High-Level Forum on Artificial Intelligence (AI)*, svoltosi in Africa dal 3 al 5 giugno 2024 sotto l'egida dell'UNESCO, che ha costituito il preludio del lancio dell'*African Consensus of Rabat on Artificial Intelligence: A Call to Action.*

Il primo, e più importante, è rappresentato dalla centralità acquisita dal principio dello *human oversight* (o della supervisione umana), da declinare anche come diritto alla autonomia della persona rispetto alla macchina e della precedenza della prima rispetto alla seconda. La precedenza della persona nella gestione e nel governo dei sistemi di intelligenza artificiale realizza, così, il superamento definitivo della visione, opposta, che le riteneva in grado di autoregolamentarsi senza incorrere in lesioni di diritti fondamentali della persona e che rendeva, quindi, recessiva l'esigenza di ricollocare al centro quest'ultima.

Lo *human oversight* costituisce la *ratio* degli interventi di regolamentazione.

I due aspetti – la riconosciuta imprescindibilità della supervisione umana e la positivizzazione di norme di legge – si tengono insieme. La consapevolezza dell'urgenza di un intervento umano è andata, cioè, di pari passo con l'abbandono di impostazioni di autoregolamentazione per dare ingresso a tentativi più o meno stringenti di normare i sistemi di intelligenza artificiale.

Sempre nella prospettiva dei profili che accomunano le regolamentazioni, attuate o potenziali, si collocano, in secondo luogo, i principi, che sviluppano i contenuti e le implicazioni dello *human oversight*.

Il riferimento è agli altri principi, c.d. diritti "nuovi" o *AI-oriented*, che si ritrovano nei testi approvati e in discussione. Si tratta del c.d. *right to know*, quale diritto della persona alla conoscenza e alla spiegabilità (*explainability*) del funzionamento dei sistemi di intelligenza artificiale; della trasparenza (*transparency*); dell'esigenza di assicurare uno sviluppo delle tecnologie di intelligenza artificiale, che sia sicuro, robusto e affidabile (*reliability*).

Terzo elemento di sintonia è costituito, infine, dallo sviluppo di strategie di valutazione del rischio e dell'impatto delle tecnologie di intelligenza artificiale, i cui primi esempi sono in fase di ultimazione da parte delle principali organizzazioni sovranazionali e di alcuni Stati.

La sintonia teorica e di principio, che si è imposta entro la dimensione globale, a livello domestico e sovranazionale, non ha, però, interessato, alcuni profili altrettanto cruciali della configurazione degli interventi di regolamentazione.

Veniamo, quindi, alle dissonanze.

In estrema sintesi, sembra di poter sostenere che si scorgono soluzioni eterogenee quanto: alla definizione di intelligenza artificiale, alla scelta dei criteri di classificazione delle nuove tecnologie e, ancora prima, alla volontà di introdurne categorie, alle modalità di valutazione dell'impatto dei sistemi di intelligenza artificiale sui diritti fondamentali.

Su tutti questi aspetti non esiste un *consensus*, con la sola eccezione dell'esperienza europea (Unione Europea e Consiglio d'Europa) e, anzi, l'evoluzione delle proposte legislative o di *policy* che si stanno moltiplicando in tutte le aree del mondo testimonia, piuttosto, una volontà di ricercare soluzioni proprie, *Country* oppure *Continent-specific*.

Difetta, così, anzitutto, unanimità con riferimento alla definizione di intelligenza artificiale. Si passa da nozioni ridotte, ma più coerenti con i portati delle scienze informatiche e che circoscrivono il concetto di intelligenza artificiale ai soli sistemi *machine* e *deep learning*, come nel caso europeo, che ripiega, come visto, sull'opzione definitoria avallata dall'OECD; a proposte definitorie alternative, quali quelle emerse nelle realtà continentali africane e asiatica; a soluzioni, che omettono di parlare di "intelligenza artificiale", preferendovi nozioni quali "sistemi automatizzati", senza distinguere tra tecnologie complesse e semplici *software*.

La dissonanza sul piano definitorio si riflette, poi, sull'ambito applicativo delle regolamentazioni in esame, che sta generando una frammentarietà che non sposa la natura *borderless* e globale del fenomeno, rischiando di pregiudicando uno scambio proficuo inter-statale di *policies* e *best practices*.

Anche le scelte di classificare le tecnologie di intelligenza artificiale in categorie, prima, e di elevare il rischio a criterio di demarcazione, poi, non superano i confini europei.

L'Unione Europea rimane, quindi, la sola organizzazione, a optare per una classificazione rigida e astratta delle tecnologie di intelligenza artificiale, laddove l'analisi comparata dimostra, viceversa, come più diffusa sia l'identificazione di sistemi di intelligenza artificiale vietati, senza distinguere, come fa l'*AI Act*, tra vietati e "ad alto rischio". In proposito, però, esiste una omogeneità sul piano delle tecnologie da vietare così come delle deroghe da introdurre al loro utilizzo. Si pensi, per fare un esempio, ai sistemi biometrici di riconoscimento facciale, il cui impiego continua ad essere generalmente ammesso se sorretto da ragioni di sicurezza e di sorveglianza pubblica.

Tornando alle divergenze, se non si concorda sulla classificazione delle tecnologie di intelligenza artificiale, lo stesso vale per la delimitazione dell'oggetto delle valutazioni di impatto. E, invero, i modelli di *risk assesment* sono, in alcuni casi, quello europeo, limitati alle sole tecnologie qualificate "ad alto rischio", in altri, estese a tutte preferendo una impostazione che non esclude che sistemi meno sofisticati siano comunque suscettibili di esporre a rischio i diritti fondamentali.

Infine, un tema su cui significative sono le distanze tra le esperienze regolatorie, è la *liability*, la ricerca, cioè, e l'individuazione del oppure dei responsabili di una violazione di uno o più diritti fondamentali imputabile al funzionamento di un sistema di intelligenza artificiale.

A questo riguardo, sono stati agli Stati degli Stati Uniti a fornire le risposte più organiche e convincenti, sul modello della legge approvata dallo Stato del Colorado. Legge che, per la prima volta, propone uno schema di responsabilità ampio, che non riposa esclusivamente sul *deployer*, affiancandovi il programmatore e, soprattutto, dettagliando gli obblighi di entrambi allo scopo di agevolare l'accertamento giudiziale dei profili di *liability*.

Oltre alle convergenze e dissonanze tra chi ha deciso di regolamentare e alla "contro-regolamentazione", cioè l'opposizione alla regolamentazione, è interessante, in chiusura, richiamare le altre tendenze che stanno segnando la c.d. *AI race*.

La prima attiene al fenomeno della *AI Countergovernance*, che ha conosciuto inveramento entro l'ordinamento canadese, a cui se ne accosta, una seconda, la *Participatory AI*, che si è, viceversa, imposta, almeno nelle intenzioni, nelle proposte avanzate dalla organizzazione degli Stati africani, degli Stati del Sud-est asiatico e, sebbene in modo meno esplicito, dalle Nazioni Unite e dall'UNESCO.

Due tendenze o approcci, che investono, come visto, il metodo e, in particolare, la scelta di "chi" consultare quando sia in discussione l'approvazione di una normativa che investa le modalità di impiego delle tecnologie di intelligenza artificiale.

La *AI Countergovernance* contesta, così, scelte di *governance*, che omettano una consultazione ampia di tutti i soggetti coinvolti. La *Participatory AI*, da parte sua, risponde a queste critiche, esprimendosi in favore di modelli di *governance* che beneficino del contributo di una cerchia vasta ed eterogenea di *stakeholders*, certo non limitabile ai rappresentanti delle industrie *tech*, come si sta verificando, ad esempio, nel Regno Unito.

Questi due modi di guardare al "come" governare e normare le tecnologie di intelligenza artificiale saranno sempre più centrali nei prossimi mesi e non si esclude che avranno una ricaduta anche per il continente europeo, in attesa della implementazione integrale dell'*AI Act*. Una circostanza, quella che ambirebbe a che fossero presi sul serio i presupposti di fondo della *AI Countergovernance* e della *Participatory AI*, che dovrebbe suggerire qualche cautela in più nel ricorrere ad approcci rigidamente *top-down*, favorendo soluzioni partecipate, come sta tentando di fare, e si vedrà con quali risultati, la Commissione europea per la stesura del suo *Code of Practice*[96].

***

Abstract: The contribution provides a comparative overview of the global approaches towards AI regulation.

Relying on domestic and supranational responses, the article argues that at least four different approaches are emerging: hard law regulation, as symbolically embraced by the EU AI Act; the AI contra-regulation, which opts for the lack of legislative regulation in the context of AI, now endorsed by the United States; the new phenomenon of the AI-counter governance, that opposes non-participatory strategies in the governance of AI, as in Canada; the preference for principles-based approaches

[96] In tema, si veda la notizia pubblicata sul sito della Commissione dell'UE, con cui è stata avviata una consultazione che ha portato sinora alla redazione di tre draft preliminari della *Code of Practice* previsto ai sensi dell'art. 95 *AI Act*. La notizia relativa all'avvio della *targeted consultation* è consultabile dal 30 aprile 2025 al seguente link: https://digital-strategy.ec.europa.eu/en/news/commission-seeks-input-clarify-rules-general-purpose-ai-models.

(e.g. the United Kingdom), where legislators refrain from detailed and strict regulations, confining their role to settling guiding principles and criteria of AI use and implementation.

The article does not omit to consider the role of international organizations, following the leading role of the Council of Europe and its Framework Convention on artificial intelligence, human rights, democracy, and the rule of law.

Ultimately, the article discusses the highly fragmented reality of AI, human rights, and the law, showing the lack of homogeneous approaches both in terms of approaches and content.

Keywords: Artificial intelligence; Regulation, Law; AI Contra-Regulation; AI Counter-Governance; Principles-based Approach; Human Rights.

***

Costanza Nardocci – Professoressa associata in Diritto costituzionale, Università degli Studi di Milano (costanza.nardocci@unimi.it)