# QUALE MODELLO SOCIALE PER L'*AI ACT* EUROPEO? NOTE MINIME SU DIRITTI E DEMOCRAZIA AL BIVIO*

di Mariavittoria Catanzariti**



«le forze economiche non debbono dominare gli uomini,
ma – come avviene per le forze naturali –
essere da loro sottomesse, guidate,
controllate nel modo più razionale,
affinché le grandi masse non ne sieno vittime»[1]

Sommario. 1. Introduzione. – 2. Una genealogia insolita: dall'etica al diritto dell'IA. – 3. L'approccio antropocentrico: da precetto etico a norma tecnico-giuridica. – 4. I rimedi del giudice europeo (all'assenza dei rimedi giurisdizionali nell'*AI Act*). – 5. Governance libera tutti. – 6. I limiti di un approccio risk-based: molto rumore per nulla. – 7. Coazione a ripetere: colonialismo regolatorio *vs* colonialismo tecnologico. – 8. Conclusioni: ripensare i presupposti.

***1. Introduzione.*** La reputazione talvolta dipende dall'affidamento ingenerato nel maggior numero di soggetti in ordine a una certa competenza. Talaltra è invece relativa al grado di affidabilità connesso al funzionamento di un sistema. Il primo modello regolatorio dell'IA *made in the EU* è certamente un buon esempio di quanto il valore della reputazione possa essere foriero di modelli culturali, anche tali da compensare il deficit tecnologico. Il marchio del buon regolatore che l'UE ha guadagnato nel corso dell'intero processo di integrazione europea gioca un ruolo primario anche nella strategia europea sui dati. Nel solco del successo post GDPR, l'UE prova a forgiare una impalcatura normativa per l'IA con aspirazioni globali in termini di ambito di applicabilità e premesse culturali locali in termini di presupposti. Il risultato, non sempre coerente, è tuttavia ambizioso a livello sistematico. Un'architettura del compromesso basata da un lato sull'idea di accettabilità del rischio e dall'altro sulla responsabilità da prodotto difettoso che concilia diverse tendenze: quella di proteggere i diritti degli utenti e quella di chiudere il cerchio delle responsabilità creando orizzontalmente alcune obbligazioni specifiche come risposta ai rischi strutturali[2]. Naturalmente la possibilità di dettare formalmente le regole del gioco e conservare nel tempo un ranking di rinomato regolatore ha il suo prezzo: che la concentrazione del capitale nelle mani dei giganti digitali

* *Sottoposto a referaggio*.
** Ricercatrice di Filosofia del Diritto – Università di Padova.
[1] Manifesto di Ventotene.
[2] Sul punto si veda G. Malgieri, *The EU AI Act: The Objectives of the Law (Article 1)*, in G. Malgieri, G.González Fuster, A. Mantelero e G. Zanfir-Fortuna (a cura di), *The EU Artificial Intelligence Act. A Thematic Commentary*, Oxford, 2026, 7-21.

non sia pregiudicata in toto, ivi compreso la clemenza con riguardo a eventuali profili sanzionatori. L'articolo tenterà di elaborare alcune chiavi di lettura critiche a partire da tale premessa scettica sul ruolo virtuoso della regolazione in assenza di pratica sociale. In particolare, la prima parte si concentrerà sul passaggio dall'etica al diritto dell'IA; la seconda analizzerà i limiti del cosiddetto approccio antropocentrico; la terza esaminerà il sistema dei rimedi giurisdizionali dell'IA Act; la quarta avrà ad oggetto il sistema di governance dell'IA basato sulla cosiddetta concezione antropocentrica dell'IA; infine, la riflessione si soffermerà sul concetto di colonialismo regolatorio a partire dall'analisi sulla portata applicativa dell'IA.



***2. Una genealogia insolita: dall'etica al diritto dell'IA.*** La modalità con cui l'Unione Europea ha inteso sviluppare il proprio approccio etico e giuridico all'intelligenza artificiale (IA) è del tutto peculiare. L'approccio europeo all'AI è stato infatti segnato dapprima dalla redazione delle *Ethics Guidelines for a Trustworthy AI* cui ha fatto seguito cinque anni dopo il Regolamento europeo (EU) 2024/1689 sull'intelligenza artificiale (cosiddetto *AI Act*)[3]. Il primo dato che emerge è dunque di sistema: il consenso etico sui valori è frutto di una *policy* preparatoria rispetto alla regolazione. Entrambi i sistemi di regole, sia quello etico sia quello giuridico, sono stati parametrati sulla base di soglie di accettabilità del rischio dei sistemi di intelligenza artificiale rispetto ai valori dell'Unione, che incarnano il limite di tenuta del sistema europeo rispetto al quale istituzioni e individui edificano la propria identità «europea». Un doppio binario, quello etico e giuridico, che duplica, almeno apparentemente i contenuti, ma che corre in parallelo, rivolgendosi a destinatari potenzialmente diversi. Le ragioni sono molteplici: anzitutto, la matrice culturale dell'approccio europeo all'IA, che riprende *latu sensu* le premesse del costituzionalismo novecentesco ispirate alla inderogabilità dei diritti della persona come assetto valoriale[4]; in secondo luogo, dal punto di vista istituzionale,

[3] Gli antecedenti storici del processo che ha portato alla redazione degli Orientamenti Etici sono stati i seguenti: la Proposta di Risoluzione del Parlamento Europeo recante raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica (Rapporto Delvaux), la Risoluzione del Parlamento Europeo recante raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica del 16 febbraio 2017; il *Committee of Experts of Internet Intermediaries* «*Algorithms and Human Rights. Study on the human rights dimensions of automated data processing techniques and possible regulatory implications*» del 2018 del Consiglio d'Europa; il vertice di Tallinn del 29 settembre 2017; la Comunicazione della Commissione Europea del 25 aprile 2018 «L'intelligenza artificiale per l'Europa» (Comunicazione COM (2018) 237) con la quale è stata lanciata la Strategia sull'Intelligenza artificiale; l'istituzione dell'High-Level Expert Group on Artificial Intelligence; la Comunicazione COM (2018) 795 «Piano coordinato sull'intelligenza artificiale»; la *European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their environment adopted by the CEPEJ durting its 31st Plenary meeting* del 3/4 dicembre 2018. Gli atti intervenuti invece tra la redazione degli Orientamenti Etici e l'adozione dell'*AI Act* sono stati i seguenti: Libro bianco sull'intelligenza artificiale (Un approccio europeo all'eccellenza e alla fiducia, 19 febbraio 2020, COM(2020) 65 final; Comunicazione della Commissione COM(2020) 67final «Plasmare il futuro digitale dell'Europa», Bruxelles, 19 febbraio 2020; Relazione della Commissione al Parlamento europeo, al Consiglio e al Comitato economico e sociale europeo sulle implicazioni dell'intelligenza artificiale, dell'internet delle cose e della robotica in materia di sicurezza e di responsabilità, Bruxelles, 19.2.2020 COM(2020) 64final; la Comunicazione della Commissione COM (2021) 118 final «*Digital Compass: the European way for the Digital Decade*» del 9 marzo 2021; infine, dopo la presentazione della Proposta di Regolamento sull'intelligenza artificiale, la Dichiarazione europea sui diritti e i principi digitali per il decennio digitale COM (2022) 28 final. Per un'analisi puntuale cfr. G. Alpa, *Quale modello normativo europeo per l'intelligenza artificiale?*, in D. Buzzelli, M. Palazzo (a cura di), *Intelligenza artificiale e diritti della persona*, Pisa, 2022, 17-38; C. Camardi (a cura di), *La via europea per l'intelligenza artificiale*, Milano, 2022, *passim*, e all'interno specificamente A. Simoncini, *Quale modello per la regolazione dell'intelligenza artificiale? L'Europa al bivio*, 239-265; C. Grieco, *Intelligenza artificiale e tutela degli utenti nel diritto dell'Unione Europea*, Napoli, 2023, *passim*.

[4] Cfr tra tutti F. Balaguer Callejón, *La costituzione dell'algoritmo*, 2023, Mondadori, *passim*; O. Pollicino, *Di cosa parliamo quando parliamo di costituzionalismo digitale?*, in *Quad. Cost.*, 3, 2023, 569-594; Id., *Costituzionalismo digitale.*

l'importanza dei valori fondamentali dell'Unione, tra i quali il rispetto della dignità umana, la libertà, la democrazia, l'uguaglianza, lo Stato di diritto e il rispetto dei diritti umani, nella lettera dell'art. 2 TUE[5]; inoltre, nella prospettiva delle ambizioni di crescita, il modello di mercato virtuoso che si intende incarnare, volto a coniugare la concorrenza e la protezione individuale, nel quale l'etica agisce a livello reputazionale e comportamentale con effetti diffusi; infine, il tentativo dell'Europa di guadagnare un ruolo di affidabile interlocutore globale sullo scenario internazionale, oscillante irrimediabilmente tra il vecchio ruolo egemonico e le nuove sfide volte alla competizione planetaria[6]. Queste ragioni informano in generale l'approccio all'innovazione che è tendenzialmente riconducibile al diritto europeo delle tecnologie informatiche e dell'IA e che mira a coniugare la realizzazione del mercato unico con la protezione dei diritti fondamentali[7].

Il compromesso europeo nutre l'ambizione di rendere praticabile l'idea di un'umanità attrezzata e ben disposta ad assumere il rischio di manipolazione ad opera delle macchine intelligenti, purché tale rischio possa essere impedito o minimizzato. Ma di quale umanità si tratta? La questione è certamente complessa perché dipende in primo luogo dalla percezione sociale e condivisa della realtà, percezione che si frammenta in singole individualità atomizzate, cui si chiede un'autonomia in certo senso automatica, immateriale, astratta, di facciata. L'autonomia è un concetto che si afferma anzitutto attraverso l'edificazione dell'io e la capacità di dissentire. Anche la capacità di agire è un concetto che giuridicamente astrae alcune facoltà ricomprese nella possibilità di dissentire[8]. Se invece la performance tecnologica dell'AI è affidata a logiche binarie che non permettono spazio di manovra, si comprende allora come la sostituzione dei diritti attraverso l'affidabilità dell'IA è in grado di mettere in sordina silenziosamente i diritti e guardare alla loro azionabilità sempre in maniera sospettosa. La performance dell'IA è infatti costruita *ex ante* secondo una logica che deve necessariamente condurre a un risultato rispetto al quale il diritto ha scarsa capacità di incidenza se posticcio. Una umanità individualistica che può al più perseguire l'obiettivo di proteggere un piccolo nucleo di prerogative, come fossero proprie soltanto e avulse dalle relazioni sociali, che se esercitate da sole e in maniera decontestualizzata e astorica non sortiscono alcun effetto modificativo rispetto al potere pressoché illimitato delle *big tech* e dunque rispetto alla realtà dei rapporti di forza. L'intelligenza artificiale viene dunque considerata dal legislatore europeo, al di là del *passepartout* formale e pacificato regolatorio, come un settore dell'innovazione nel quale è auspicabile che l'individuo eserciti il suo controllo in forme piuttosto rarefatte rispetto allo spazio di diritti, in quanto deliberatamente volto a realizzare condizioni di restrizione della libertà sotto diverse forme. Si tratta semplicemente di stabilire le forme di limitazione della libertà nella maniera più indolore possibile. Antropocentrismo dunque inteso come possibilità di scegliere alternativamente rispetto a un percorso segnato, in virtù dello status morale unico e inalienabile di primato in

*Pensare la democrazia al tempo dell'IA*, Bolgona, 2025; H.W. Micklitz, O. Pollicino, A. Reichman, A. Simoncini, G. Sartor, G. De Gregorio (a cura di), *Constitutional Challenges in the Algotìrithmic Society*, Cambridge, 2021, *passim*.

[5] Sul ruolo dei valori europei si veda M. Almada, G. De Gregorio, *Product Safety and Fundamental Rights Regulation*, in G. Malgieri, G. González Fuster, A. Mantelero, G. Zanfir-Fortuna (a cura di), *The EU Artificial Intelligence Act. A Thematic Commentary*, Oxford, 2026, 61-74; A. von Bogdandy, *The Emergence of European Society through Public Law. A Hegelian and Anti- Schmittian Approach*, Oxford, 2024, 96-113.

[6] P. Rossi, *Il ruolo dell'Europa: origini e declino di un'asimmetria*, in A. Ruberti (a cura di), *Europa a confronto. Innovazione, tecnologia, società*, Roma-Bari, 391-437.

[7] Sul punto si veda A. Pirozzoli, *The Human-Centric Perspective in the Regulation of Artificial Intelligence*, in *European Papers*, 9, 2024, 105-116; M. Almada, N. Petit, *The EU AI Act: between the rock of product safety and the hard place of fundamental rights*, in *Common market law review*, 1, 2025, 85-120.

[8] G. Alpa, *Status e capacità: la costruzione giuridica delle differenze individuali*, Roma-Bari, 1993, 63.

campo civile, politico, economico e sociale riservato all'essere umano[9]. Tuttavia, a ben vedere, asserire lo *status* morale dell'individuo non fornisce alcuna indicazione sulla possibilità di esercitare diritti, che invero nel modello regolatorio europeo della IA appaiono molto limitati. Occorre infatti notare che i diritti individuali richiamati dall'*AI Act* siano minimi e si riducano al diritto a ricevere dall'utilizzatore una spiegazione o di presentare reclamo all'autorità di vigilanza del mercato ma non richiamano i corrispondenti rimedi giurisdizionali presenti per la violazione di altri diritti, come ad esempio per la protezione dei dati personali, per la tutela del principio di buon andamento dell'amministrazione, per il divieto di discriminazione. Ciò costituisce un *vulnus* significativo, poiché, indipendentemente dal fatto che un sistema possa determinare effetti negativi e che tali effetti siano preventivamente testati rispetto all'immissione del sistema sul mercato, in realtà non esiste nel testo normativo la previsione di una modalità umana di interazione con i sistemi di intelligenza artificiale, i cui effetti fisici e psicofisici possono determinarsi anche molto tempo dopo l'utilizzo e in virtù di logiche diverse che nulla hanno a che fare con il principio di causalità del danno. È evidente che rispetto a questo effetto il diritto è nudo, e la mutazione del contesto di riferimento determina dipendenza rispetto al potere economico e tecnologico ed eteronomia implicita, che agisce in ogni caso a livello istituzionale.

Basti pensare all'efficientamento di modelli predittivi applicabili in svariati ambiti nei quali è alto il rischio di arrecare pregiudizio alla democrazia, alla *rule of law* e ai diritti, come l'amministrazione della giustizia, lo sviluppo dei processi democratici, il controllo delle frontiere, l'educazione, tutti ambiti questi coperti dall'allegato III all'*AI Act* che enumera, assieme ai sistemi IA considerati ad alto rischio dall'art. 6, c. 1-3, *AI Act*, anche un elenco tassativo di contesti applicativi di sistemi ad alto rischio tipizzati. Per sistemi ad alto rischio si intendono tanto i sistemi di IA destinati a essere utilizzati come componenti di sicurezza di un prodotto, o i sistemi di IA che sono essi stessi prodotti, quanto i prodotti, il cui componente di sicurezza è il sistema di IA, o i sistemi di IA stesso in quanto prodotti, che sono soggetti a una valutazione della conformità da parte di terzi ai fini dell'immissione sul mercato o della messa in servizio di tali prodotti sul mercato europeo. Sono altresì considerati ad alto rischio i sistemi che effettuano profilazione di individui e quelli contenuti nell'allegato III dell'*AI Act*, relativi ad esempio alla biometria, all'istruzione, al lavoro, ai servizi pubblici essenziali, all'amministrazione della giustizia e ai processi democratici, alle attività di contrasto e di immigrazione, asilo e controllo delle frontiere, quando presentano un rischio significativo di danno per la salute, la sicurezza o i diritti fondamentali delle persone fisiche, anche se non direttamente volti a influenzare materialmente il risultato del processo decisionale.

In primo luogo, l'assuefazione rispetto a una idea di sostituibilità dell'umano non tecnologicamente neutrale e servente rispetto alla concentrazione del capitale nelle mani dei colossi digitali, forgia un modello di intelligenza artificiale funzionale alla produzione di profitto e di tecniche di sorveglianza, riducendo il più possibile il dissenso, senza contrappesi istituzionali volti a creare spazi di critica e di partecipazione politica[10]. In secondo luogo, data la pervasività dei sistemi AI nella vita quotidiana, non è chiaro quale tipo di individuo sia votato all'esercizio del controllo, se l'individuo colto, dotato di particolari capacità cognitive, competenze e strumenti culturali e materiali, o l'individuo standardizzato, profilato, ipnotizzato dai banner pubblicitari, targetizzato o offuscato nella propria capacità critica. Inoltre, appare poco probabile che quello stesso individuo possa formare autonomamente

[9] Orientamenti etici per una intelligenza artificiale affidabile, § 38 e *AI Act*, § 1.

[10] J. Cupać, M. Sienknecht, *Regulate against the machine: how the EU mitigates AI harm to democracy*, in *Democratization*, 5, 2024, 1067-1090.

una opinione se tale opinione non è il frutto di un libero scambio dialettico con altri soggetti[11]. In sostanza, a fronte di un modello regolatorio dettagliato come quello europeo, non è dato comprendere che tipo di processo di formazione dell'opinione pubblica possa determinarsi con e all'interno di un ambiente in cui si utilizza l'intelligenza artificiale, soprattutto considerando che lo sviluppo dell'IA è completamente riposto nelle mani delle *big tech*[12] e dell'assenza di competitività con attori alternativi.

***3. L'approccio antropocentrico: da precetto etico a norma tecnico-giuridica.*** Gli orientamenti etici dell'*AI HLEG* riferiscono il requisito dell'intervento e della sorveglianza umani allo sviluppo e all'utilizzo dei «sistemi di IA come strumenti al servizio delle persone, nel rispetto della dignità umana e dell'autonomia personale, e funzionano in modo da poter essere adeguatamente controllati e sorvegliati dagli esseri umani» (§ 27). Emerge chiaramente sin da subito il carattere ambivalente dell'intelligenza artificiale che, se da un lato può produrre un'«ampia gamma di benefici a livello economico, ambientale e sociale nell'intero spettro delle attività industriali e sociali» (Cons. 4 *AI Act*), dall'altro può «comportare rischi e pregiudicare gli interessi pubblici e i diritti fondamentali tutelati dal diritto dell'Unione» (Cons. 5 *AI Act*). In termini pratici, un approccio antropocentrico implica che si possa permettere a una persona fisica di intervenire, interrompere o rettificare il risultato di un algoritmo nel caso di errore, discriminazione o decisione non spiegabile; che la supervisione umana sia progettata e garantita nell'intero ciclo di vita del sistema e che ciò sia documentato nelle specifiche tecniche; che l'assenza di meccanismi di controllo umano costituisce una violazione sanzionata dall'*AI Act*. Tale architettura, infatti, si fonda su tre pilastri: centralità dell'uomo e dei diritti fondamentali (art. 1 e considerando 1-5); sorveglianza umana e *accountability* (artt. 14 e 29); gestione dei rischi e trasparenza algoritmica (artt. 9-13).

Le basi giuridiche dell'*AI Act*, essendo contenute nello stesso articolo 1 che manifesta espressamente l'adesione al modello antropocentrico, perseguono una chiara finalità: se da un lato è formalmente prescritta e adottata la procedura legislativa ordinaria, dall'altro è evidente l'intenzione di armonizzare le disposizioni legislative, regolamentari e amministrative degli Stati membri che incidono sull'instaurazione e sul funzionamento del mercato interno[13]. Il modello di una intelligenza artificiale antropocentrica e affidabile deve essere in sostanza compatibile con la garanzia di «un livello elevato di protezione della salute, della sicurezza e dei diritti fondamentali sanciti dalla Carta dei diritti fondamentali dell'Unione europea, compresi la democrazia, lo Stato di diritto e la protezione dell'ambiente, contro gli effetti nocivi dei sistemi di IA nell'Unione, e promuovendo l'innovazione» (art.1 *AI Act*). La chiave di volta di tale compatibilità è da rinvenire nella nozione di rischio, definito come «la combinazione della probabilità del verificarsi di un danno e la gravità del danno stesso» (art. 3, n. 2, *AI Act*) sulla base della quale i sistemi di IA sono classificati come a rischio minimo, alto o inaccettabile. Tra le opzioni regolatorie disponibili antecedentemente alla proposta della Commissione vi erano le seguenti: predisporre uno strumento legislativo dell'UE che istituisse un sistema di etichettatura volontario; mettere a punto un approccio settoriale ad hoc; adottare uno strumento legislativo orizzontale dell'UE che seguisse un approccio proporzionato e basato sul rischio; migliorare tale ultima opzione aggiungendo la possibilità



[11] J. Habermas, *Storia e critica dell'opinione pubblica*, Roma-Bari, 2006 (ed.or. Strukturwander der Öffentlichkeit. Untersuchungen zu einer Kategorie der bürgerlichen Gesellschaft, Frankfurt am Main, 1962), 33 ss.
[12] D. Guarascio, *Imperialismo Digitale. Economia e guerra ai tempi delle piattaforme e dell'IA*, Roma-Bari, 2026, 69.
[13] T. Brown, D. E. Harasimiuc, *Legal, Regulatory and Ethical Non-binary Choices of the AI Act*, Londra, 2025, 166-178.

dell'adozione di codici di condotta per i sistemi di IA non ad alto rischio (opzione prescelta); infine, adottare uno strumento legislativo orizzontale dell'UE che stabilisca requisiti obbligatori per tutti i sistemi di IA, indipendentemente dal loro rischio. Si è optato alla fine per il penultimo modello, al fine di creare una differenziazione tra sistemi ad alto rischio e sistemi a rischio minimo.

La prospettiva antropocentrica basata sul rischio si riflette inoltre nella possibilità di sorveglianza umana che si fonda sulla competenza, sulla autorità e sulla formazione in relazione alle seguenti attività (art. 14 *AI Act*): a) comprendere correttamente le capacità e i limiti pertinenti del sistema di IA ed essere in grado di monitorarne debitamente il funzionamento, anche al fine di individuare e affrontare anomalie, disfunzioni e prestazioni inattese; b) restare consapevoli della possibile tendenza a fare automaticamente affidamento o a fare eccessivo affidamento sull'output prodotto da un sistema di IA, in particolare in relazione ai sistemi di IA ad alto rischio utilizzati per fornire informazioni o raccomandazioni per le decisioni che devono essere prese da persone fisiche; c) interpretare correttamente l'output del sistema di IA, tenendo conto ad esempio degli strumenti e dei metodi di interpretazione disponibili; d) decidere, in qualsiasi situazione particolare, di non usare il sistema di IA ad alto rischio o altrimenti di ignorarne, annullarne o ribaltarne l'output; e) intervenire sul funzionamento del sistema di IA o interromperne il funzionamento mediante un pulsante di «arresto» o una procedura analoga che consenta al sistema di arrestarsi in condizioni di sicurezza. Per i sistemi di identificazione biometrica, l'*AI Act* prevede inoltre un doppio meccanismo di verificazione con almeno due persone fisiche.



Un ulteriore aspetto relativo all'approccio antropocentrico riguarda i processi decisionali *data-driven* e in particolare i processi decisionali automatizzati. La computabilità e la tracciabilità dei dati e delle interazioni sociali che derivano dal loro uso – si pensi ai sistemi di intelligenza artificiale per finalità generali – ci pongono l'interrogativo di quali siano le opportunità di mutamento sociale ed emancipazione nell'ambito di modelli predittivi che mirano a dar forma a un ordine sociale sulla base della razionalità algoritmica. Tale tipo di razionalità non esigerebbe – al contrario del potere legale razionale – obbedienza *ex post* bensì *ex ante*, in quanto la totale obbedienza, sotto forma di esecuzione procedurale, è presupposto essenziale della performance tecnologia. Se ci si vuole servire di una certa tecnologia, è necessario sottostare pedissequamente al suo diktat senza possibilità di diversa interpretazione del comando, altrimenti non si può fruire di quella tecnologia. Questo fenomeno si rende ancor più evidente nella propria insostenibilità sociale quando si tratta di tecnologie basate sull'uso di dati passati per inferire previsioni sul futuro che hanno un impatto sulla libertà di autodeterminazione individuale.

Rispetto a tali pratiche appare singolare che l'*AI Act* non preveda, come anticipato, rimedi giurisdizionali specifici se non il diritto al reclamo nei confronti dell'autorità di vigilanza del mercato per violazione delle norme del Regolamento ai sensi dell'art. 85 *AI Act* e il diritto in capo a qualsiasi persona interessata ritenga di aver subito un effetto giuridico o un impatto significativo tale da far percepire un effetto negativo sulla salute, sulla sicurezza e sui diritti fondamentali a opera dell'utilizzo del sistema AI ovvero riceva una decisione basata sull'output di un sistema ad alto rischio di cui all'allegato III dell'*AI Act*, a ottenere dall'utilizzatore la spiegazione sul ruolo giocato dal sistema IA nella procedura decisionale e sui principali elementi della decisione adottata. Tale ultimo diritto non era contenuto nella proposta della Commissione ma è stato inserito a opera di un emendamento del Parlamento. Vale la pena riflettere sulla portata applicativa estremamente limitata della disposizione in oggetto. Intanto si tratta di un diritto che si riferisce al solo utilizzatore dei sistemi IA per finalità professionali e al solo soggetto interessato ai sensi del GDPR, cioè a una persona fisica identificata o identificabile. Inoltre, il diritto alla spiegazione previsto dall'*AI Act* è

escluso se è già previsto da altra normativa nazionale o europea, come ad esempio dal GDPR con riguardo agli obblighi informativi e alla decisione umana, e dal *Digital Services Act* con riguardo ai parametri utilizzati per le pratiche di targeting pubblicitario. L'unico elemento di maggiore tutela rispetto ad altri strumenti come il GDPR consiste nel fatto che la spiegazione ai sensi dell'*AI Act* prescinde dal danno; dunque, può essere tanto una spiegazione generalizzata quanto una spiegazione individualizzata circa la finalità del processo decisionale, il tipo di approccio e di tecnologia adottati, le categorie di dati utilizzati e il loro impatto nella determinazione dei risultati, l'individuazione dei dati inferenziali.



***4. I rimedi del giudice europeo (all'assenza di rimedi giurisdizionali dell'AI Act).*** Nel contesto dei rimedi giurisdizionali riguardanti i processi decisionali automatizzati, come spesso accade, il diritto pretorio europeo ha assunto un ruolo molto più rilevante, anche se limitato a una interpretazione evolutiva del GDPR, rispetto alla scarna filigrana delle disposizioni sui mezzi di ricorso previste dall'*AI Act*, e in futuro potrebbe probabilmente instaurare alcune connessioni tra il diritto alla spiegazione di cui all'art. 86 *AI Act* e il diritto alla decisione umana di cui all'art. 22 GDPR[14]. Ciò è accaduto perché l'*AI Act* annovera tra i mezzi di ricorso soltanto il diritto di presentare reclamo alla competente autorità di vigilanza del mercato per violazioni delle disposizioni dell'*AI Act* e di ottenere dal *deployer* spiegazioni chiare e significative sul ruolo del sistema di IA nella procedura decisionale e sui principali elementi della decisione adottata[15]. Quest'ultimo diritto (cosiddetto diritto alla spiegazione), tuttavia, ha un'applicazione residuale rispetto ad analoghi diritti previsti dal diritto dell'Unione. Di fatto, dunque, l'*AI Act* non prevede rimedi giurisdizionali specifici, fatti salvi naturalmente altri ricorsi amministrativi o giurisdizionali. Per questa ragione la tessitura più robusta del GDPR è stata adattata ai processi decisionali automatizzati seppur per il tramite della protezione dei dati personali.

Il 7 dicembre 2023, la Corte di giustizia dell'UE (CGUE) ha emesso la sua prima sentenza sul processo decisionale individuale automatizzato nella causa C-634/21 (caso *Schufa*). La Corte ha stabilito che qualsiasi tipo di valutazione automatizzata è vietata se ha un impatto significativo sulla vita individuale. La sentenza riguardava la SCHUFA, una grande agenzia di credito privata tedesca che valuta le persone in base al merito creditizio e assegna loro un punteggio. La CGUE ha stabilito che il *credit scoring* deve essere interpretato come una decisione automatizzata in quanto produce una decisione automatizzata. Ha quindi applicato il criterio dell'effetto giuridico che, sebbene determinato dall'attività di due parti e in diverse fasi del trattamento, era principalmente legato al comportamento degli istituti di credito, ossia alla considerazione dei punteggi di rischio algoritmici durante le trattative.

Il giudice del rinvio ha chiesto se l'articolo 22, paragrafo 1, del GDPR debba essere interpretato come un «processo decisionale automatizzato relativo a persone fisiche», ai sensi di tale disposizione, per una società che fornisce informazioni commerciali che calcola automaticamente un tasso di probabilità basato su dati personali relativi a una persona e riguardanti la capacità di tale persona di onorare gli impegni di pagamento in futuro, laddove la conclusione, l'esecuzione o la cessazione di un rapporto contrattuale con tale persona da parte di un terzo, al quale tale tasso di probabilità viene comunicato, dipenda in modo decisivo da tale tasso di probabilità.

[14] In tal senso, appare molto apprezzabile lo studio di E. Falletti, C. Gallesi, *Credit scoring judicial review between the Court of Justice of the European Union and comparative case Law*, in *Rivista di diritto dei media*, 3, 2024, 92-116.

[15] Cfr. sul punto M. Margieska, Remedies and Penalties (Articles 85, 87, 99-101), in in G. Malgieri, G.González Fuster, A. Mantelero, G. Zanfir-Fortuna (a cura di) *The EU Artificial Intelligence Act. A Thematic Commentary*, Oxford, 2026, 533-550.

La CGUE ha rinviato la causa al tribunale nazionale (tedesco) per stabilire se (i) nel caso in questione, i punteggi di rischio hanno svolto un ruolo decisivo nella negazione del credito e, in caso affermativo, se (ii) il *credit risk scoring* si qualifica come un'eccezione legittima al divieto generale di affidarsi a decisioni automatizzate e, in caso affermativo, se (iii) SCHUFA, in quanto agenzia di *credit scoring*, soddisfa i requisiti stabiliti dal GDPR per le decisioni automatizzate (legittime). Secondo la sentenza, lo scoring di SCHUFA viola il GDPR se i clienti di SCHUFA, come le banche, gli attribuiscono un ruolo «decisivo» nelle loro decisioni contrattuali. Per la CGUE, il calcolo delle probabilità può essere considerato un processo decisionale automatizzato ai sensi dell'articolo 22 del GDPR, se sono soddisfatte tre condizioni. In primo luogo, la decisione deve essere «basata esclusivamente su un trattamento automatizzato, compresa la profilazione» (il calcolo automatizzato di un tasso di probabilità basato su dati personali e relativo alla capacità della persona di onorare un prestito in futuro è da considerarsi un'attività di profilazione); in secondo luogo, è necessario produrre effetti giuridici che riguardino l'interessato o che incidano sulla sua persona in modo altrettanto significativo; in terzo luogo, il processo decisionale si basa in modo decisivo sul tasso di probabilità calcolato. Se tutto ciò è vero, allora l'articolo 22, che conferisce all'interessato il «diritto di non essere sottoposto a una decisione basata unicamente su un trattamento automatizzato, compresa la profilazione, è violato e tale violazione non deve essere fatta valere individualmente da tale persona. Se il calcolo del tasso di probabilità è considerato solo un atto preparatorio e l'azione intrapresa dal terzo può, se mai, essere qualificata come una «decisione», vi è il rischio di eludere l'articolo 22, che conferisce all'interessato il diritto di non essere sottoposto a una decisione basata esclusivamente su un trattamento automatizzato, compresa la profilazione. La CGUE ha quindi stabilito che il processo decisionale automatizzato basato su un calcolo probabilistico della produzione di un effetto negativo sulle persone (rifiuto di accesso al sistema creditizio) è una tecnica di profilazione e costituisce una violazione dell'articolo 22 del GDPR. Una pronuncia ancor più recente della CJEU del 27 febbraio 2025 nel caso C-203/22 CK contro *Magistrat der Stadt Wien* si pone su una scia simile agli *Schufa Cases* allargando ulteriormente l'orizzonte a una dimensione di protezione ex ante. Recupera cioè i diritti informativi di cui all'art. 15 GDPR quale garanzia funzionale della completa tutela del diritto alla decisione umana di cui all'art. 22 GDPR. E difatti la Corte di Giustizia, rispondendo a sei questioni pregiudiziali, afferma, da un lato, che il concetto di diritto di informazione può ricomprendere il diritto dell'interessato, in un contesto di trattamento automatizzato di dati personali come la profilazione, di pretendere dal titolare del trattamento «informazioni significative sulla logica utilizzata», cioè sulla procedura e sui principi concretamente applicati per utilizzare, con mezzi automatizzati, i dati personali relativi a tale interessato al fine di ottenerne un risultato determinato, come un profilo di solvibilità, mediante informazioni pertinenti e in forma concisa, trasparente, comprensibile e facilmente accessibile. Dall'altro, la Corte sancisce l'obbligo in capo al titolare del trattamento che fornisca informazioni all'interessato contenenti dati personali di terzi o dati di terzi protetti da segreto commerciale, di comunicare tali informazioni asseritamente protette all'autorità di controllo o al giudice competenti, cui spetta ponderare i diritti e gli interessi in gioco al fine di determinare la portata del diritto di accesso dell'interessato previsto all'articolo 15 di tale regolamento.



***5. Governance libera tutti.*** Un approccio basato sul rischio ha come presupposto implicito quello di riconoscerlo. Corre osservare che una società prevedibile è in primo luogo per definizione una società connotata da elementi di immutabilità nella quale i soggetti, le forme istituzionali e i processi decisionali sono frutto di una precisa progettazione delegata ad algoritmi e modelli IA sviluppati da poteri economici che costituiscono oligopoli. Essi, infatti, si spartiscono i settori rilevanti dei mercati digitali lungo tutta la filiera produttiva, dalla

produzione dei semiconduttori allo sviluppo dei software e dei modelli di intelligenza artificiale generativa, e attraverso l'analisi di una mole gigantesca di dati e la loro conseguente interpretazione unilaterale e riducibilità a fattori quantitativi, sono in grado di manipolare la realtà e offrire letture sempre più unilaterali e meno complesse[16]. Rispetto a tali pratiche, il diritto alla spiegazione avrebbe lo scopo di evitare la tautologica evidenza del dato, facendo emergere come l'uso di un sistema AI possa avere effetti trasformativi rispetto ai soggetti e ai contesti.

Nella *governance* europea dell'IA basata sul rischio, l'esigenza di non comprimere il mercato si dota di anticorpi istituzionali che riguardano le pratiche ad alto rischio, rispetto alle quali gli attori privati sono obbligati a mettere in atto alcune misure di gestione del rischio prima di affacciarsi sul mercato europeo, nonché di gestione di qualità dei sistemi di IA, a fornire le istruzioni per l'uso agli utilizzatori, oltre alla documentazione tecnica idonea a rendere le autorità nazionali competenti responsabili per la valutazione, la designazione e la notifica degli organismi di valutazione della conformità e del monitoraggio dei sistemi di IA con il diritto europeo. La *governance* del rischio comprende inoltre anche misure che assicurino la responsabilità lungo la catena del valore dell'IA, riguardanti tutti i soggetti coinvolti dalla progettazione alla utilizzazione, nonché la valutazione d'impatto dei sistemi IA sui diritti fondamentali. In particolare, il controllo dei dati e il controllo delle infrastrutture attraverso le quali i dati possono essere utilizzati rappresenta il perno assiologico sul quale l'*AI Act* ripone massima fiducia[17]. Dal punto di vista operativo, sotto il primo profilo, la governance dei dati è tradotta in una norma prevista dall'*AI Act* (art. 10) che riguarda alcune pratiche relative ai set di dati di addestramento nei sistemi ad alto rischio, in ordine a: le scelte progettuali pertinenti; i processi di raccolta dei dati e l'origine dei dati, nonché la finalità originaria della raccolta nel caso di dati personali; le operazioni di trattamento pertinenti ai fini della preparazione dei dati, quali annotazione, etichettatura, pulizia, aggiornamento, arricchimento e aggregazione; la formulazione di ipotesi, in particolare per quanto riguarda le informazioni che si presume che i dati misurino e rappresentino; una valutazione della disponibilità, della quantità e dell'adeguatezza dei set di dati necessari; un esame atto a valutare le possibili distorsioni suscettibili di incidere sulla salute e sulla sicurezza delle persone, di avere un impatto negativo sui diritti fondamentali o di comportare discriminazioni vietate dal diritto dell'Unione, specie laddove gli output di dati influenzano gli input per operazioni future; le misure adeguate per individuare, prevenire e attenuare le possibili distorsioni; l'individuazione di lacune o carenze pertinenti nei dati tali da pregiudicare il rispetto del Regolamento e il modo in cui tali lacune e carenze possono essere colmate. Tale *governance* si fonda per l'appunto sull'affidabilità degli utilizzatori e delle certificazioni riguardo ai set di addestramento. Tuttavia, la pubblicazione delle *Guidelines* a opera della Commissione Europea nel corso del 2025 aventi a oggetto tre diversi aspetti - pratiche vietate, portata degli obblighi dei provider per l'uso dei sistemi di AI per finalità generali e definizione di sistemi AI ha reso più agevole l'interpretazione pratica dell'*AI Act*[18] – è sintomatico dell'esigenza di interpretazione uniforme dal lato degli operatori, alla luce dell'entrata in vigore graduale degli obblighi derivanti dall'IA per i fornitori di sistemi con finalità generali e dei poteri di esecuzione della Commissione. Unitamente a tali orientamenti,



[16] A. Supiot, *Governance by numbers. The Making of a Legal Model of Allegiance*, Londra, 2017, 78.
[17] M. Maas, *Architectures of Global AI Governance. From Technological Change to Human Choice*, Oxford, 52, 113.
[18] Orientamenti della Commissione relativi alle pratiche di intelligenza artificiale vietate ai sensi del regolamento (UE) 2024/1689 (regolamento sull'IA), Bruxelles, 29.7.2025 C(2025) 5052 final; Orientamenti della Commissione relativi alla definizione dei sistemi di IA stabiliti dal Regolamento (EU) 2024/1689 (*AI Act*), Brussels, 29.7.2025 C(2025) 5053 final; Orientamenti della Commissione relativi all'ambito di applicazione degli obblighi relativi ai modelli di IA per uso generale stabiliti dal regolamento (UE) 2024/1689 (*AI Act*).

è stato adottato anche un Codice di buone pratiche per i fornitori di sistemi IA con finalità generali che si basa sul principio di presunzione di conformità[19]. Sebbene, dunque, l'adesione a tale codice sia volontaria, essa è funzionale a dimostrare la conformità all'*AI Act*. Ai fini definitori, ad esempio, per sistema IA si intende un sistema basato su macchine in cui il sistema di IA deve essere guidato da calcoli e basato su operazioni meccaniche, il che coinvolge sia componenti hardware sia componenti software per consentire il funzionamento dei sistemi di IA e copre i sistemi di calcolo quantistico. Tale sistema deve inoltre essere autonomo, nel senso di essere in grado di sviluppare inferenze indipendentemente dall'intervento umano. Il sistema deve inoltre poter essere adattabile, cioè avere capacità di autoapprendimento nella modifica del proprio comportamento durante l'uso. Attraverso tecniche di elaborazione dei dati, come l'apprendimento automatico (apprendimento supervisionato, non supervisionato e rinforzo) e il ragionamento basato sulla logica e sulla conoscenza, il sistema deve avere capacità inferenziale, cioè essere in grado di dedurre, dai risultati che riceve, come generare risultati. Questi risultati possono differenziarsi tra previsioni, contenuti e raccomandazioni, e per poter essere considerati risultati dell'IA devono poter influenzare ambienti fisici o virtuali. Le previsioni sono stime relative a un valore sconosciuto (output) basate su valori noti forniti al sistema (input) che richiedono il minimo intervento umano. I contenuti si riferiscono alla generazione di nuovi materiali da parte di un sistema di IA. Le raccomandazioni sono suggerimenti relativi ad azioni, prodotti o servizi specifici rivolti agli utenti in base alle loro preferenze, ai loro comportamenti o ai dati da loro inseriti. Le decisioni si riferiscono alle conclusioni o alle scelte effettuate dal sistema di IA. Infine, un sistema IA deve essere in grado di interagire con l'ambiente nel quale è implementato.

***6. I limiti di un approccio risk-based: molto rumore per nulla.*** Riguardo al limite che la regolazione europea si autoimpone, quello delle pratiche vietate dall'art. 5 *AI Act*, occorre svolgere qualche riflessione di metodo. L'art. 5 del Regolamento include un catalogo di pratiche vietate in quanto parametrate a un rischio inaccettabile e si rivolge, rispetto all'ampia gamma di soggetti previsti dall'ambito di applicazione materiale del Regolamento (fornitori, *deployer*, importatori, distributori e fabbricanti di prodotti), soltanto a fornitori e utilizzatori che possono essere sia attori pubblici sia attori privati[20]. Prevedendo una gradualità dell'assolutezza del divieto (senza o con eccezioni), esso comprende tra le pratiche vietate quelle che utilizzano tecniche subliminali visive, uditive, tattili , cioè senza che il soggetto sia consapevole, o manipolative o ingannevoli, o che sfruttano le vulnerabilità di una persona fisica o di uno specifico gruppo di persone, dovute all'età, alla disabilità o a una specifica situazione sociale o economica tali da provocare un danno significativo alla persona, o che determinino il *social scoring*, cioè un punteggio sociale alle persone in merito ai comportamenti individuali o collettivi in contesti non collegati ai contesti in cui i dati sono stati originariamente raccolti o che sia ingiustificato e sproporzionato rispetto alla entità del comportamento, pratica vietata sia nel settore pubblico sia nel settore privato se produttiva di danno; sistemi che prevedono il rischio di commissione di reati basati esclusivamente su profilazione preventiva; sistemi finalizzati a creare o incrementare banche dati di

[19] Cfr R. Romano, D. Amram, *Codes of Conduct and EC Guidelines (Articles 56, 95-98)*, in G. Malgieri, G.González Fuster, A. Mantelero, G. Zanfir-Fortuna (a cura di), *The EU Artificial Intelligence Act. A Thematic Commentary*, Oxford, 2026, 553-563.

[20] M.L. Rebrean, S. De Conca, *The Prohibitions Related to Human Vulnerabilities and Manioulation (Articles 5(1) and 5(1)(b))*, in G. Malgieri, G.González Fuster, A. Mantelero e G. Zanfir-Fortuna (a cura di), *The EU Artificial Intelligence Act. A Thematic Commentary*, Oxford, 2026, 77-104; P. Voigt, N. Hullen, *The EU AI Act Answers to Frequently Asked Questions*, Cham, 2026, 37.

riconoscimento facciale mediante *scraping*, cioè estrazione di informazioni liberamente accessibili per finalità aggregative; sistemi che inferiscano emozioni sui luoghi di lavoro o istituti di istruzione tranne se consentite per finalità mediche o di sicurezza; sistemi di categorizzazione biometrica con finalità discriminatoria in ordine a razza, opinioni politiche, appartenenza sindacale, convinzioni religiose o filosofiche, vita sessuale o orientamento sessuale, tranne che per attività di contrasto; sistemi di identificazione biometrica remota «in tempo reale» in spazi accessibili al pubblico per finalità di contrasto tranne che per sottrazione, tratta o sfruttamento sessuale di esseri umani, prevenzione di minaccia alla vita o atto terroristico e localizzazione o identificazione di una persona sospettata di un reato che preveda la pena detentiva per un massimo di almeno quattro anni[21]. Per far sì che tali pratiche siano considerate vietate, è necessario che siano integrati quattro presupposti: la pratica deve costituire l'«immissione sul mercato», la «messa in servizio» o l'«uso» di un sistema di IA; il sistema di IA deve utilizzare tecniche subliminali (che agiscono senza che una persona ne sia consapevole), volutamente manipolative o ingannevoli; lo scopo o l'effetto delle tecniche utilizzate dal sistema di IA deve essere quello di distorcere materialmente il comportamento di una persona o di un gruppo di persone. La distorsione deve risultare da un pregiudizio considerevole della capacità di prendere una decisione informata, tale da indurre una persona a prendere una decisione che non avrebbero altrimenti preso; la distorsione del comportamento deve produrre, o ragionevolmente poter produrre, un danno significativo a tale persona, a un'altra persona o a un gruppo di persone[22]. Emerge chiaramente come l'inaccettabilità del rischio sia stata costruita in astratto in relazione a una tipologia di pratiche vietate per le quali tuttavia, affinché esse integrino il divieto, è necessario che in concreto vi sia un nesso causale tra distorsione del comportamento per effetto di tali pratiche e danno causato – fisico, psicologico o economico – alla persona.

La selezione delle pratiche considerate a rischio inaccettabile ha chiaramente l'intento di evitare la diffusione di sistemi lesivi dell'autonomia individuale. Un'idea di autonomia individuale intesa come capacità di orientare scelte e azioni indipendentemente dall'esercizio distorsivo del potere economico e tecnologico. A ben vedere, tuttavia, il concetto di autonomia individuale nel contesto tecnologico è relativo, in assenza di alternative a modelli



[21] Negli Orientamenti della Commissione relativi alle pratiche di intelligenza artificiale vietate ai sensi del regolamento (UE) 2024/1689 del 29 luglio 2025, emanati dalla Commissione ai sensi dell'art. 96 *AI Act*, non vincolanti, sono considerate pratiche vietate ad esempio quelle che utilizzano tecniche subliminali che agiscono senza che una persona ne sia consapevole o tecniche volutamente manipolative o ingannevoli aventi lo scopo o l'effetto di distorcere il comportamento, che provochino o possano ragionevolmente provocare un danno significativo; che sfruttano le vulnerabilità dovute all'età, alla disabilità o a una specifica situazione sociale o economica, con l'obiettivo o l'effetto di distorcere il comportamento, che provochino o possano ragionevolmente provocare un danno significativo; che attraverso l'attribuzione di un punteggio sociale determinino un trattamento pregiudizievole o sfavorevole quando i dati provengono da contesti sociali non collegati o quando tale trattamento è ingiustificato o sproporzionato rispetto al comportamento sociale; che valutano o prevedono il rischio che le persone commettano un reato unicamente sulla base della profilazione o dei tratti e delle caratteristiche della personalità; tranne quando tali sistemi sono utilizzati a sostegno di una valutazione umana basata su fatti oggettivi e verificabili direttamente connessi a un'attività criminosa; che creano o ampliano le banche dati di riconoscimento facciale mediante *scraping* non mirato di immagini facciali da internet o da filmati di telecamere a circuito chiuso; che inferiscono le emozioni sul luogo di lavoro o negli istituti di istruzione, tranne che per motivi medici e di sicurezza; che classificano le persone sulla base dei loro dati biometrici per trarre deduzioni o inferenze in merito a razza, opinioni politiche, appartenenza sindacale, convinzioni religiose o filosofiche o orientamento sessuale; fatta eccezione per l'etichettatura o il filtraggio di set di dati biometrici acquisiti legalmente, anche nel settore delle attività di contrasto. Cfr R. J. Neuwirth, *Prohibited artificial intelligence practices in the proposed EU artificial intelligence act (AIA)*, in *Computer law & security review*, 48, 2023, 105798.

[22] Orientamenti della Commissione relativi alle pratiche di intelligenza artificiale vietate, cit., p. 21.

monopolistici che impongono attraverso gli stessi servizi digitali di larga fruizione una visione del mondo predeterminata alla quale è necessario implicitamente aderire per poterne usufruire. In secondo luogo, si considera altamente lesiva della dignità della persona e della libertà qualsiasi pratica discriminatoria che arrechi un pregiudizio anche potenziale al singolo in quanto individuo o facente parte di un gruppo sociale[23]. A ben guardare tuttavia, la selezione dei fattori di vulnerabilità limitati soltanto alla età, alla disabilità o a una specifica situazione sociale o economica non è spiegabile se non nell'ottica di immaginare che tali fattori siano funzionali allo sviluppo della personalità più di altri nell'ottica del legislatore. Circostanza che in punto di diritto è del tutto incomprensibile se posta a confronto invece con i fattori discriminatori come la razza, le opinioni politiche, l'appartenenza sindacale, le convinzioni religiose o filosofiche, la vita sessuale o l'orientamento sessuale che però sono posti al riparo soltanto dei sistemi di categorizzazione biometrica e non delle altre pratiche vietate. Non si comprende, in particolare, se vulnerabilità e discriminazione siano considerate in un'ottica comparativa dal punto di vista qualitativo ovvero quantitativo. Ad esempio, in base alla lettera del regolamento, se un sistema di IA non di categorizzazione biometrica realizzasse un effetto discriminatorio non produrrebbe una pratica vietata così come invece un sistema di sfruttamento delle vulnerabilità sarebbe lecito se non sfruttasse nello specifico fattori come età, condizione sociale o economica, disabilità bensì altri fattori di vulnerabilità inducendo all'adozione di comportamenti non autonomi e pregiudizievoli. Il catalogo delle pratiche vietate che si applicano tanto ai sistemi con finalità prestabilita tanto a quelli con finalità generali, nonostante gli sforzi esplicativi contenuti nelle *Guidelines* della Commissione Europea, tradiscono in realtà una scarsa chiarezza nelle intenzioni

Allo stesso tempo, desta non poche perplessità la disposizione che non vieta la pratica di *social scoring tout court*, come scelta di sistema, ma soltanto qualora essa integri un danno (tecnicamente «trattamento pregiudizievole o sfavorevole») per persone o gruppi in contesti sociali scollegati dal contesto di raccolta dei dati ovvero ingiustificato o sproporzionato rispetto al loro comportamento sociale o alla sua gravità[24]. In fin dei conti, dunque, il divieto di *social scoring* è piuttosto un limite di facciata rispetto a un modello, quello cinese ad esempio, che ha tutt'altra ispirazione, basandosi sul concetto di fiducia totale nei confronti dello stato[25]. La trasposizione di tale divieto nella cultura giuridica occidentale rischia travisamenti evidenti se si realizza ricorrendo all'uso di categorie proprie del diritto moderno e non è neanche filtrato attraverso categorie di adattamento[26].

Una perplessità ancora maggiore è invece suscitata dalle pratiche ad alto rischio incluse nell'allegato III all'*AI Act*, che sono tuttavia considerate tali soltanto quando si configuri un rischio significativo per la salute, la sicurezza o i diritti fondamentali idoneo a influenzare

---

[23] F. Rimoli, *Sulla dignità "in senso oggettivo" come argomento giuridico*, in *Diritto & Società*, 4, 2024, 661-687, 685: «"la diade concettuale dignità/libertà si pone come luogo centrale dell'autorappresentazione individuale e fondamento del principio di autodeterminazione del singolo" … "Sono dunque i diritti, così concepiti, riconosciuti e tutelati, a generare dignità e libertà quali spazi dell'autonomia individuale, e non l'opposto". In sostanza, aderendo a una concezione luhmaniana, è la protezione dei diritti a generare dignità e libertà e non l'opposto».

[24] In particolare, la contrarietà al social scoring è stata accentuata negli slogan più che nella lettera del legislatore, al fine di evidenziare come l'Europa dei diritti fondamentali si collochi su un piano diametralmente opposto rispetto alla Repubblica Popolare Cinese. A. Vigorito, *Art. 5 (comma 1 c)*, in A. Mantelero, G. Resta, G. M. Riccio (a cura di) *Intelligenza Artificiale. Commentario*, Milano, 2025, 143-151.

[25] M. S. Erie, T. Streinz, *The Beijing Effect: China's 'Digital Silk Road' as Transnational Data Governance*, in *J. Int'l L. & Pol.*, 1, 2021, 54.

[26] T. W. Adorno, *Individuum und Staat*, 1951, in *La società degli individui*, 3, 2000, 9 (tr. it. I. Testa). Si tratta di un testo inedito, pubblicato per la prima volta in R. Tiedemann (a cura di), *Gesammelte Schriften*, Vermischte Schriften, 20, 1, Frankfurt a. M. 1986, 287-292.

materialmente il risultato del processo decisionale. L'elenco dei sistemi IA contenuti nell'Allegato III possono essere oggetto di modifica e di riduzione a opera dei decreti delegati della Commissione Europea. Ad oggi include sistemi di identificazione e categorizzazione biometrica e riconoscimento delle emozioni; sistemi inerenti a infrastrutture critiche; sistemi relativi a profili specifici dell'educazione e della formazione professionale, come l'accesso, l'ammissione e il reclutamento, nonché la valutazione dell'apprendimento, del livello di istruzione e il monitoraggio di comportamenti vietati; sistemi riguardanti il reclutamento nei rapporti di lavoro e la valutazione della performance lavorativa; sistemi di accesso a prestazioni private o pubbliche, come accesso al credito, alle assicurazioni, alla sanità e alla previdenza; sistemi utilizzati per attività di contrasto per valutare il rischio di essere vittima di reato, l'affidabilità degli elementi probatori in sede di indagine o di perseguimento, il rischio di recidiva, nonché la profilazione in corso di indagine, accertamento e perseguimento dei reati; sistemi relativi alla migrazione, asilo e gestione delle frontiere utilizzati dalle autorità competenti anche nella valutazione dei rischi alla sicurezza, e nella gestione delle domande di asilo, visto o permesso di soggiorno per valutarne l'ammissibilità o per individuare, identificare o riconoscere persone fisiche; infine, sistemi utilizzati nell'ambito dell'amministrazione della giustizia o dei processi democratici, per assistere l'autorità giudiziaria nella ricerca, nell'interpretazione dei fatti o nell'applicazione della legge a un fatto concreto, ossia per influenzare l'esito di una votazione. Da questo nutrito elenco emerge chiaramente come molteplici settori della vita democratica quali il processo, l'amministrazione della giustizia, le elezioni, il controllo dei flussi migratori, nonché il nucleo essenziale dei diritti sociali, politici e civili, come lavoro, istruzione, voto, libertà personale e identità, siano completamente sdoganati rispetto all'uso di sistemi di intelligenza artificiale che evidentemente costituiscono un alto rischio, benché anche soltanto potenziale, per la capacità di alterazione dei processi sottostanti. Come si può, a questo punto, immaginare che tali pratiche non soltanto siano sempre consentite ma siano considerate ad alto rischio, richiedendo dunque misure idonee a mitigarlo da parte delle imprese, solo se comportano un rischio significativo alla salute, alla sicurezza e ai diritti fondamentali, senza che ciò determini un cambiamento radicale nei processi democratici?[27] L'art. 6, c. 2 e 3, *AI Act* include infatti tra i sistemi ad alto rischio quei sistemi compresi nell'Allegato III che presentano un rischio significativo di danno per la salute, la sicurezza o i diritti fondamentali delle persone fisiche, menzionando espressamente il rischio di influenzare materialmente il risultato del processo decisionale e in ogni caso i sistemi che effettuino profilazione di persone fisiche. Non sono considerati ad alto rischio, tuttavia, questi stessi sistemi qualora essi siano approntati per eseguire un compito procedurale limitato, a migliorare il risultato di un'attività umana precedentemente compiuta, a rilevare schemi decisionali o deviazioni da schemi decisionali precedenti e non finalizzati a sostituire o influenzare la valutazione umana precedentemente completata senza un'adeguata revisione umana, infine a eseguire un compito preparatorio per una valutazione pertinente ai fini dei casi d'uso elencati nell'allegato III. Ancora una volta, dunque, la regolazione non fa che attestare una situazione di fatto, senza tuttavia costituire un reale limite all'indebolimento della democrazia e dei diritti. Difatti ciò che colpisce è la lettura dell'allegato III in combinato con il considerando 62 che si riferisce ai «sistemi di IA destinati a essere utilizzati per influenzare l'esito di elezioni o referendum o il comportamento di voto delle persone fisiche nell'esercizio del loro voto alle elezioni o ai referendum dovrebbero essere classificati come sistemi di IA ad alto rischio, ad



[27] G. Resta, C. Frattone, *Classification of High-Risk AI Systems and Its Update (Articles 6, 7, 50, 71, Annex I and Annex III)*, in G. Malgieri, G. González Fuster, A. Mantelero, G. Zanfir-Fortuna (a cura di), *The EU Artificial Intelligence Act. A Thematic Commentary*, Oxford, 2026, 181-212.

eccezione dei sistemi di IA ai cui output le persone fisiche non sono direttamente esposte, come gli strumenti utilizzati per organizzare, ottimizzare e strutturare le campagne politiche da un punto di vista amministrativo e logistico». Ciò conferma ancora una volta, da un lato, l'approccio meramente individualistico dell'*AI Act*, volto a tutelare l'integrità del voto del singolo individuo e non invece il processo comune di formazione dell'opinione pubblica, dall'altro, un'attitudine del tutto formale al confronto con fenomeni complessi, che esclude dal novero delle pratiche manipolative realizzate attraverso i social media quelle che si prefiggono lo scopo di gestire campagne elettorali dal punto di vista logistico e amministrativo, tenendo conto che però, tra esse, sono sempre considerate ad alto rischio le pratiche di profilazione volte alla scelta dei luoghi per i comizi elettorali[28]. Tale traiettoria individua una curva ascendente nella previsione volta a limitare il *deep fake*, che obbliga gli utilizzatori a rendere noto che il contenuto è stato generato o manipolato artificialmente (art. 50, c. 4). Nei processi democratici il problema di fondo è, però, ciò che Habermas definisce il problema della rappresentanza pubblica del dominio[29]. Se tale rappresentanza si spezzetta, si affievolisce, si disperde, la voce della democrazia, cioè la pubblica opinione volta all'esercizio della critica, risulta completamente neutralizzata da un potere che stenta a farsi riconoscere.



***7. Coazione a ripetere: colonialismo regolatorio vs colonialismo tecnologico.*** Sulla scia della efficacia extraterritoriale del GDPR, anche l'*AI Act* evidenzia un aspetto evidente relativo all'ambito di applicazione: la promessa di non applicabilità dell'*AI Act* nei confronti di chi esporta sistemi di intelligenza artificiale fuori dall'Unione Europea. Mentre infatti si assiste a una amplissima efficacia extraterritoriale dell'*IA Act* nei confronti di tutti i sistemi destinati al mercato europeo, alla sua applicazione sfugge l'export europeo dei modelli IA. Questo è particolarmente rilevante nei settori comunque sottratti all'ambito di applicazione materiale dell'*AI Act* quali la sicurezza nazionale, la difesa e l'uso militare, nonché nei confronti di autorità pubbliche di paesi terzi e organizzazioni internazionali. L'*AI Act*, come anticipato, essendo uno strumento a efficacia orizzontale, si applica a una pluralità di soggetti, tra i quali i «fornitori», i loro «rappresentanti autorizzati», i «*deployers*», gli «importatori» e i «distributori», i «fabbricanti di prodotti» che incorporano dispositivi di i.a., le «persone interessate» ai sensi del GDPR. Inoltre i casi di portata territoriale del regolamento europeo includono i seguenti casi: immissione nel mercato o messa in servizio nel mercato UE da parte di utilizzatori anche se non stabiliti o ubicati all'interno dell'UE di sistemi di IA o modelli di IA per finalità generali nell'Unione; utilizzatori stabiliti o ubicati all'interno dell'UE; fornitori e utilizzatori nell'Unione stabiliti o ubicati in un paese terzo, qualora l'*output* prodotto sia utilizzato nell'UE; importatori e distributori di sistemi AI; fabbricanti di prodotti che immettono sul mercato o mettono in servizio un sistema di IA insieme al loro prodotto e con il loro nome o marchio; rappresentanti dei fornitori non stabiliti in UE; persone interessate che si trovano in UE.

Mentre l'immissione sul mercato, definita come «fornitura di un sistema di IA o di un modello di IA per finalità generali per la distribuzione o l'uso sul mercato dell'Unione nel corso di un'attività commerciale, a titolo oneroso o gratuito», art. 3, n.10, riguarda importatori e distributori di sistemi di i.a. (lett. d), fabbricanti di prodotti che immettono sul mercato o mettono in servizio un sistema di IA insieme al loro prodotto e con il loro nome o marchio

[28] Per un'analisi puntuale cfr R. Rubio, F. F. Alvim, *Artificial intelligence and electoral campaigns*, Hoboken, 2026.
[29] J. Habermas, *Cultura e critica. Riflessioni sul concetto di partecipazione politica e altri saggi*, Torino, 1980, 54.

(lett. e), rappresentanti autorizzati di fornitori, non stabiliti nell'Unione (lett. *f*), la messa in servizio, intesa come «fornitura di un sistema di IA direttamente al *deployer* per il primo uso o per uso proprio nell'Unione per la finalità prevista», coinvolge utilizzatori, cioè soggetti pubblici e privati, ma non nell'esercizio di un'attività personale. L'output prodotto in Europa si riferisce al risultato prodotto dal sistema AI, che l'*AI Act* distingue dunque dall'utilizzo del sistema IA, e che può essere privo di limitazioni temporali e causali rispetto all'utilizzo del sistema IA, e riguarda i fornitori e utilizzatori stabiliti o situati in un paese terzo[30]. Occorre considerare che mentre l'*AI Act* non si applica ai sistemi che vengono addestrati e sviluppati in Europa ma che non sono utilizzati in Europa, al contrario si applica a qualsiasi fornitore che immetta sul mercato o metta in servizio sistemi IA nell'Unione Europea. Tra i profili maggiormente critici, emergono quelli contemplati nei considerando 21 e 22 dell'*AI Act*, allorché si legge, nel secondo, che l'«operatore stabilito nell'Unione che appalta alcuni servizi a un operatore stabilito in un paese terzo in relazione a un'attività che deve essere svolta da un sistema di IA che sarebbe classificato ad alto rischio. In tali circostanze il sistema di IA utilizzato dall'operatore in un paese terzo potrebbe trattare dati raccolti nell'Unione e da lì trasferiti nel rispetto della legge, e fornire all'operatore appaltante nell'Unione l'output di tale sistema di IA risultante da tale trattamento, senza che tale sistema di IA sia immesso sul mercato, messo in servizio o utilizzato nell'Unione». Ciò comporta un ampliamento considerevole della portata applicativa finanche alla fase dell'addestramento dei dati in Europa se utilizzati da un soggetto appaltatore di un paese terzo che fornisca l'output a un soggetto appaltante europeo. Ancor più ampia appare la previsione relativa a sistemi non ancora utilizzati nell'UE, nella misura in cui l'output prodotto da tali sistemi sia destinato a essere utilizzato nell'Unione. Nel considerando 21 si legge infatti: «Al fine di garantire condizioni di parità e una protezione efficace dei diritti e delle libertà delle persone in tutta l'Unione, è opportuno che le regole stabilite dal presente regolamento si applichino ai fornitori di sistemi di IA in modo non discriminatorio, a prescindere dal fatto che siano stabiliti nell'Unione o in un paese terzo, e ai *deployer* dei sistemi di IA stabiliti nell'Unione».
A fronte di una così ampia portata territoriale, che rasenta i limiti dell'indeterminatezza, non si può fare a meno di notare come essa sia difficilmente coniugabile con la mera autonomia funzionale di una regolazione avulsa dal controllo delle infrastrutture digitali. Su che cosa poggia dunque la fiducia nel potere della regolazione che privilegia la libera concorrenza a scapito della spesa pubblica e degli investimenti infrastrutturali? Il sodalizio tra i colossi digitali e i settori militari sugella l'idea che l'*export* sia più redditizio dell'*import*, specie se lontano dai riflettori. Il paradosso dell'effetto pregiudizievole per le imprese europee e estere che privilegiano il mercato europeo per le proprie attività rispetto a imprese che destinano i loro prodotti di sistemi IA a mercati extraeuropei tradisce a ben vedere una vocazione ondivaga e flessibile rispetto alla regolazione, che si presta a una sua applicazione intermittente se pur costruita attraverso una sofisticata architettura[31].

***8. Conclusioni: ripensare i presupposti.*** La traiettoria tracciata dal modello europeo di regolazione dell'IA rappresenta indubbiamente un notevole sforzo che tenta di tenere insieme innovazione e diritti. Tuttavia, se si torna con la mente al Manifesto di Ventotene, non si può non rammentare un presagio che con l'accresciuto e incondizionato potere in capo alle *big tech* è diventato una sconcertante realtà. Il Manifesto prefigurava infatti i rischi



[30] G. Resta, *Commento all'Art. 2 (commi 1-4), Reg. 13.6.2024, n. 2024/1689/UE*, in A. Mantelero, G. Resta, G. M. Riccio (a cura di), *Intelligenza Artificiale. Commentario*, Milano, 2025, 11-22.
[31] M. Almada, A. Radu, *The Brussels side-effect: how the AI Act can reduce the global reach of EU policy*, in *German Law Journal*, 4, 2024, 646-663.

della privatizzazione e della delega completa al privato nel processo di industrializzazione, che avrebbe irrimediabilmente creato un connubio indissolubile tra capitale e potere pubblico, servente rispetto alla politica a scapito dei deboli: «Non si possono più lasciare ai privati le imprese che, svolgendo un'attività necessariamente monopolistica, sono in condizioni di sfruttare la massa dei consumatori; ad esempio le industrie elettriche, le imprese che si vogliono mantenere in vita per ragioni di interesse collettivo ma che, per reggersi, hanno bisogno di dazi protettivi, sussidi, ordinazioni di favore ecc. (l'esempio più notevole di questo tipo d'industria sono finora in Italia le siderurgiche); e le imprese che per la grandezza dei capitali investiti e il numero degli operai occupati, o per l'importanza del settore che dominano, possono ricattare gli organi dello stato, imponendo la politica per loro più vantaggiosa (es.: industrie minerarie, grandi istituti bancari, grandi armamenti). È questo il campo in cui si dovrà procedere senz'altro a nazionalizzazioni su scala vastissima, senza alcun riguardo per i diritti acquisiti»[32].



Dinanzi a questa profezia, un tipo di regolazione basata in larga parte su sistemi di autocertificazione e oneri burocratici in capo ai *provider*[33], non riesce da sola a neutralizzare il potere delle *big tech* sul mercato europeo, ma si rivela uno specchietto per le allodole in termini di efficacia, oltre che un onere disincentivante per le start up che provino ad affacciarsi sul mercato europeo. Gli utenti, di contro, non appaiono per nulla gratificati dall'approccio antropocentrico, non potendo di fatto incidere in maniera determinante sull'alterazione dei processi democratici a opera dell'IA. Il diritto europeo dell'IA non può pretendere dunque di sorvolare sui diritti individuali e sul sistema di contrappesi istituzionali, e al contempo rivendicare una vocazione quasi universale, salvo doverlo immaginare come formula vuota rivolta a un beneficiario indifferenziato, cioè a un individuo astoricizzato e immutabile. Mentre si sa, i diritti non possono che essere sempre il frutto di lotte e rivendicazioni, e dunque situati storicamente[34]. L'etica dell'IA, al di là di un'ornata operazione di *make-up* ideologico, ha prodotto un effetto importante di *double bind* normativo, favorendo la deregolamentazione. Lo schiacciamento della regolazione su una logica estrattiva del dato ha favorito un modello di diffidenza o nella migliore delle ipotesi di indifferenza nei confronti della creatività individuale. Sia l'individuo sia le istituzioni rappresentano molto di più rispetto alla loro capacità difensiva o a un mero ruolo di spettatori. Il mito europeo della culla del diritto necessita dunque di ripensare i suoi presupposti in maniera sostanziale se non vuole subire soluzioni ibride che non hanno impatto a livello globale, come dimostra la recente proposta di semplificazione degli obblighi *Digital Omnibus*, lontano dal clamore degli antagonismi e ripartendo dal carattere necessario e mai superfluo dei diritti. Quando i diritti necessitano di declamare il loro soggetto come fosse un visconte dimezzato, pacificato dall'IA affidabile, ecco che diventano superflui, e dunque, non sono credibili.

***Abstract.*** L'adozione dell'*AI Act* rinsalda il collaudato ruolo della UE di *first mover* regolatorio digitale, inaugurato con il Regolamento Generale sulla protezione dei dati personali (GDPR) e adattato nel corso di tutta la *Data Strategy*, con l'obiettivo di costruire un modello giuridico ispirato alla compatibilità tra libertà economiche e diritti fondamentali.
I capisaldi dell'*AI Act* sono essenzialmente quattro: l'approccio antropocentrico fondato sulla

[32] *Per un'Europa libera e unita. Manifesto di Ventotene*, 2017, https://www.senato.it/application/xmanager/projects/leg18/file/repository/relazioni/libreria/novita/XVII/Per_unEuropa_libera_e_unita_Ventotene6.763_KB.pdf, 63-65.
[33] P. Lee, *Obligations of Providers of High-Risk AI Systems (Articles 8, 9, 16, 20, 21, 22)*, in G. Malgieri, G. González Fuster, A. Mantelero, G. Zanfir-Fortuna (a cura di), *The EU Artificial Intelligence Act. A Thematic Commentary*, Oxford, 2026, 213-223.
[34] N. Bobbio, *L'età dei diritti*, Torino, 1990, 20-23.

dignità della persona e sul controllo umano dell'intero ciclo di vita dei sistemi; l'approccio *risk-based* basato sull'accettabilità del rischio dei sistemi di intelligenza artificiale; l'ampia portata applicativa materiale e territoriale del regolamento europeo; la tassonomia dei sistemi ad alto rischio inclusi quelli di cui all'allegato III la cui modifica è di competenza della Commissione Europea.
Il contributo si soffermerà su questi elementi mettendone a fuoco tematicamente i tratti distintivi, con l'intento precipuo di riflettere su come l'insieme di alcuni fattori contribuisca a costruire il modello regolatorio europeo della AI e quali siano le implicazioni principali. Ove letta in chiave causale, la concomitanza di alcuni fattori dà luogo a una riflessione inevitabile sul rapporto tra intelligenza artificiale e democrazia. L'architettura giuridica complessiva europea dell'intelligenza artificiale, infatti, lascia adito a non poche criticità, che il contributo tenterà di contestualizzare all'interno di un preciso impianto concettuale, avente a oggetto i punti di forza e quelli di inciampo del complesso rapporto tra diritti e innovazione. L'angolo visuale dal quale muove il contributo indaga se e in quale misura la combinazione tra approccio antropocentrico, approccio basato sul rischio, portata applicativa dell'*AI Act* e tassonomia dei sistemi ad alto rischio sia in grado di comporre un quadro di garanzie sufficienti a livello istituzionale e intersoggettivo tale da preservare i principi democratici.



**Abstract.** With the adoption of the AI Act, the EU is consolidating its proven role as a digital regulatory first mover, which began with the General Data Protection Regulation (GDPR) and has been adapted throughout the Data Strategy, with the aim of building a legal model inspired by the compatibility between economic freedoms and fundamental rights. There are essentially four cornerstones to the AI Act: the anthropocentric approach based on human dignity and human control over the entire life cycle of systems; the risk-based approach based on the acceptability of risk in artificial intelligence systems; the broad material and territorial scope of application of the European regulation; the taxonomy of high-risk systems, including those referred to in Annex III, which the European Commission is responsible for amending.
This contribution will focus on these elements, highlighting their distinctive features, with the primary aim of reflecting on how a combination of factors contributes to the construction of the European regulatory model for AI and what the main implications are. When read in a causal light, the group of some features inevitably leads to a reflection on the relationship between artificial intelligence and democracy. The overall European legal architecture of artificial intelligence, in fact, leaves room for a number of critical issues, which the contribution will attempt to contextualise within a precise conceptual framework, focusing on the strengths and stumbling blocks of the complex relationship between rights and innovation. The perspective from which this contribution proceeds investigates whether and to what extent the combination of an anthropocentric approach, a risk-based approach, the scope of application of the AI Act and the taxonomy of high-risk systems can provide a framework of sufficient institutional and intersubjective guarantees to preserve democratic principles.

***Parole chiave.*** Rischio – Approccio antropocentrico – Diritto&Etica dell'IA – Governance dell'IA – Diritti fondamentali.

***Key words.*** Risk – Anthropocentric approach – Law&Ethics of AI – AI governance – Fundamental Rights.