La regolazione dei sistemi ad alto rischio nell'Artificial Intelligence Act: considerazioni d'insieme

1. I nuclei problematici della disciplina

Uno sguardo d'insieme rivolto alla disciplina dei sistemi ad alto rischio non permette di soffermarsi sul dettaglio del testo legislativo, ma impone di concentrarsi su alcune scelte architetturali. Le due principali questioni che vorrei esaminare riguardano, da una parte, la stessa individuazione dei sistemi ad alto rischio; da un'altra parte, l'opzione verso una disciplina di misure preventive, consistenti in requisiti essenziali di sicurezza, la cui attuazione è rimessa al metodo della coregolazione con il coinvolgimento degli organismi di standardizzazione europei.

L'identificazione dei sistemi ad alto rischio è stata operata attraverso una selezione articolata e complessa[1], che richiede di essere vagliata nel merito al fine di evidenziare il perimetro operativo delle regole e riconoscere eventuali applicazioni che, per quanto non prive di criticità, ne risultano escluse. All'esito di questa ricerca si tratterà di trovare risposte giuridiche alternative o complementari per quei casi, avulsi dalla copertura offerta dalla definizione dell'art. 6, che comunque sollecitano il ricorso a qualche forma di tutela.

Quanto alla soluzione di disciplina che vede affiancati requisiti cd. di alto livello e standard tecnici, occorre esaminare criticamente la trasposizione di questo modello dall'ambito del controllo di sicurezza che investe molti settori produttivi connotati da rischi specifici, di natura prevalentemente materiale, all'area nevralgica delle applicazioni di intelligenza artificiale, caratterizzate viceversa dalla diffusione in contesti molto diversi e da rischi di tipo immateriale.

2. La natura duale dei sistemi ad alto rischio tra continuità e discontinuità

Com'è noto, l'identificazione dei sistemi ad alto rischio è avvenuta con la definizione posta dall'art. 6 dell'AI Act, da leggere in combinazione con gli Allegati I e III. Tale nozione mette insieme due classi molto diverse di prodotti: quelli già sottoposti a una legislazione europea armonizzata e, all'interno di questa, vincolati a una verifica della conformità da parte di un organismo terzo e indipendente (art. 6, comma 1, lett. a) e b)), e quella dei sistemi digitali autosufficienti impiegati in otto diverse aree critiche elencate nell'Allegato III.

L'eterogeneità di tali categorie è evidente: si riflette anzitutto nel metodo che si è scelto per designarle, con un rinvio ad altri corpi normativi, nel primo caso, e l'enumerazione tassativa nel secondo. Ma soprattutto risponde a una diversa fisionomia delle due ipotesi, contrassegnate, ancorché in termini non assoluti, dall'integrazione in un dispositivo tangibile in un caso, e dall'avere natura digitale nell'altro. Questa considerazione unitaria si pone in continuità con l'approccio che risale all'avvio della Strategia per un'IA europea, dove non si distingue tra intelligenza incorporata ovvero che abbia natura di puro software, privo di un supporto materiale, e si usa la medesima espressione come un "termine ombrello", obliterando per vero talune differenze nella tipologia di rischi che essa pone, e di danni che ne costituiscono la materializzazione in caso di incidenti o fallimenti.

[1] La determinazione delle categorie di rischio, cui corrispondono i diversi regimi, non è fatta secondo criteri puramente tecnici, bensì "è scelta che esprime una discrezionalità politica": A. Oddenino, *Intelligenza artificiale e tutela dei diritti fondamentali: alcune notazioni critiche sulla recente Proposta di Regolamento della UE, con particolare riferimento all'approccio basato sul rischio e al pericolo di discriminazione algoritmica*, in *Intelligenza artificiale e diritto: una rivoluzione?*, vol. 1, *Diritti fondamentali, dati personali e regolazione*, a cura di A. Pajno, F. Donati e A. Perrucci, Bologna, il Mulino, 2022, p. 198.

Non è quindi sorprendente che le due classi siano accomunate sotto un'unica disciplina, in simmetria con la logica, tutto sommato persuasiva, che il target della regolazione debba essere il modulo autonomo e dinamico che guida i sistemi e innesca le azioni, le quali a loro volta avranno un impatto nel mondo fisico o nell'ambiente digitale, a prescindere dalla dimensione materiale o virtuale del sistema medesimo.

Salvo ritornare nel prossimo paragrafo sulla nozione di sistemi ad alto rischio e, soprattutto, su cosa stia dentro e cosa fuori rispetto a questa cornice, merita sottolineare un ultimo aspetto. La definizione generale dei sistemi *stand alone*, in termini astratti comunque piuttosto comprensiva, viene disarticolata attraverso un intrico di deroghe, che potrebbero finire per alterare il rapporto regola/eccezione. Non sarebbero infatti ad alto rischio quei sistemi che, pur rientrando nella nozione generale, in concreto non pongono rischi per i diritti fondamentali perché svolgono un compito procedurale circoscritto, hanno un ruolo solo preparatorio, servono per migliorare i risultati di un'attività umana svolta precedentemente, valgono semplicemente a rilevare schemi o deviazioni da schemi di decisione già assunti senza sostituirli (art. 6, comma 3). L'accertamento di queste circostanze, peraltro, è fatto dallo stesso fornitore che lo assevera autonomamente. Esiste insomma un reticolo di pesi e contrappesi, che aprono a margini di discrezionalità piuttosto ampi nella messa in opera della disciplina da parte di fornitori e utenti, e poi nel controllo amministrativo e giurisdizionale.

Può dirsi, in conclusione, che la selezione dei sistemi ad alto rischio e i contorni mobili della categoria cerchino di sostanziare un approccio europeo non improntato in maniera rigida al principio di precauzione, ossia alla ricerca della minimizzazione dei rischi attraverso cautele di ordine massimo, bensì diretto a ottimizzare il controllo dei rischi compatibilmente con lo sviluppo delle libertà economiche e l'assegnazione di sufficienti spazi di manovra ai vari attori dell'innovazione[2]. Fin dall costruzione della piramide del rischio e nella selezione delle ipotesi destinate a comporne i gradini, il regolamento si propone quale disciplina orientata essenzialmente alla promozione del mercato.

3. La qualificazione e la disciplina dei sistemi: lacune, sovrapposizioni …

La stessa nozione di sistema ad alto rischio designa altresì il perimetro applicativo di un cospicuo fascio di previsioni e segna pertanto il crinale tra una regolazione dettagliata e una semi-assenza di regole. Occorre dunque capire le ragioni che hanno portato a inserire nella relativa categoria alcune ipotesi e, per converso, a lasciarne fuori altre, provando a tracciare un bilancio sulla sua latitudine e sulla qualità sufficientemente inclusiva[3].

Al netto delle questioni che riguardano la prima classe di sistemi ad alto rischio (ad esempio, il paradosso che contrassegna il caso dei veicoli autonomi[4]), più interessante, ai fini della valutazione che stiamo compiendo, è la categoria dei sistemi cd. *stand alone*, destinati ad essere adoperati in otto aree chiave, a loro volta elencate nell'Allegato III. Questa lista – che va dal controllo di infrastrutture critiche (n. 2), alle istituzioni scolastiche e universitarie (n. 3), al contesto lavorativo (n. 4), all'accesso e la fruizione di servizi essenziali, pubblici o privati (n. 5) – potrà essere modificata con un processo semplificato e secondo la metodologia di *risk assessment* prevista dall'art. 7, ma soltanto nel senso

[2] G. De Gregorio e P. Dunn, *The European Risk-Based Approaches: Connecting Constitutional Dots in the Digital Age*, in "Common Market Law Review", 2022, 59(2), p. 477 s.

[3] Per un giudizio nei termini della *under-* od *over-inclusiveness* v. già P. Hacker, *The European AI liability directives – Critique of a half-hearted approach and lessons for the future*, in "Computer Law and Security Review", 2023, 51, pp. 9 ss.

[4] Per un'analisi più diffusa sia consentito rinviare a E. Palmerini, *La governance dei sistemi ad alto rischio nell'Artificial Intelligence Act: uno sguardo panoramico*, in "Foro it.", 2025, V, c. 36.

di aggiungere o rimuovere casi di uso all'interno dei settori già individuati, che sono viceversa stabili. Si tratta dunque di chiedersi se sia sufficientemente inclusiva ed equilibrata, anche in rapporto alla necessità di non creare un carico eccessivo in termini di *compliance* per gli operatori[5]; e, soprattutto, di capire se vi siano sovrapposizioni con altri presidi di regolazione che possano essere gestite in maniera sinergica, ovvero se si prospettino ambiti scoperti da colmare in via interpretativa.

Dai documenti di accompagnamento al testo legislativo si evince che la selezione degli "*use-cases*" è stata fatta applicando i criteri del tipo, della severità e della probabilità del rischio di danno, a partire dallo screening di una platea molto ampia di ipotesi tratte da rapporti elaborati da diversi gruppi di esperti, da documenti di istituzioni nazionali e internazionali, dalla letteratura accademica o da analisi di organizzazioni non governative, dalla consultazione pubblica, infine, a cui era stato esposto il Libro Bianco sull'IA del 2020. La scelta di escludere talune applicazioni è dipesa poi, per un verso, dalla metodologia seguita, cosicché sono rimaste fuori dall'elenco quelle che non sembravano presentare rischi di danno alla salute, alla sicurezza o ad altri diritti fondamentali, oppure i cui rischi non avevano una dimensione importante; per un altro verso, dalla circostanza che alcune tecniche, pur potenzialmente altamente rischiose, sono governate da altri plessi normativi.

Una rassegna delle ipotesi menzionate nei lavori preparatori e molte altre su cui si sta formando una vasta letteratura mostrano come una casistica assai varia sia rimasta estranea alla classificazione dei sistemi ad alto rischio. Si tratta dei sistemi di raccomandazione integrati nei social network, nei videogames, nei servizi di news online, che hanno l'effetto di mediare l'accesso all'informazione, all'intrattenimento, agli acquisti; degli assistenti personali incorporati nei più comuni dispositivi elettronici, anche indossabili, che gestiscono le interazioni con l'ambiente digitale; delle numerosissime applicazioni che forniscono consigli su salute e benessere, consulenza di natura legale o finanziaria; dei meccanismi usati dalle piattaforme e dai motori di ricerca per organizzare, filtrare e presentare i contenuti agli utenti. Non trovano quindi diretto riconoscimento nelle fattispecie di alto rischio congegnate dall'AI Act un'ampia congerie di tecnologie cd. della persuasione, ossia tutte quelle funzionalità, abilitate da algoritmi estremamente diffusi, rivolte a sollecitare l'attenzione e a massimizzare la permanenza degli utenti su un certo applicativo fino a provocare vere e proprie forme di dipendenza; a indirizzare i comportamenti nella sfera dei consumi e sollecitare finanche forme di acquisto compulsivo; a controllare la conoscenza a cui si viene esposti, potenzialmente impoverendo il pluralismo dell'informazione.

A queste apparenti lacune è possibile invero reagire in due modi: per un verso, sfruttando le potenzialità di qualificazioni plurime dello stesso sistema offerte dallo stesso AI Act; per un altro verso, valorizzando le sinergie con altri ordini giuridici.

Sotto il primo profilo, occorre notare come alcuni sistemi siano suscettibili di essere inquadrati in diverse classi di rischio, cui corrisponde un regime differente. Un esempio interessante, a questo riguardo, è quello dei sistemi di riconoscimento delle emozioni[6]. Diverse compagnie attive nel settore dell'high-tech hanno progettato e commercializzato applicazioni con potenzialità di impiego svariate, dalla selezione dei candidati a una posizione lavorativa, al *neuromarketing*, al campo medico, dove

[5] Esprime la giusta preoccupazione che un approccio "oneroso e indifferenziato" per tutte le applicazioni graverà in maniera sproprorzionata sulle piccole imprese e le start-up G. Finocchiaro, *Intelligenza artificiale. Quali regole?*, Bologna, il Mulino, 2024, pp. 121 s.

[6] Questi sistemi stanno da tempo suscitando l'interesse della letteratura civilistica: cfr. E.M. Incutti, *Sistemi di riconoscimento delle emozioni e ruolo dell'autonomia privata: linee evolutive di un umanesimo digitale*, in "Giustizia civile", 2022, pp. 151 ss.; E. Tuccari, Neuromarketing: *un'asistemica disciplina … oltre il consenso?*, in "Persona e mercato", 2024, pp. 511 ss.; R. Montinaro, *Riconoscimento delle emozioni e marketing personalizzato*, ivi, pp. 847 ss.; T. De Mari Casareto Dal Verme, *Artificial Intelligence, Neuroscience and Emotional Data. What Role for Private Autonomy in the Digital Market?*, in "Erasmus Law Review", 2023, n. 3.

sono implicate specialmente nella diagnosi di stati mentali patologici[7]. La tecnologia che è alla base di questi prodotti invero è piuttosto contesa, poiché si dubita che abbia basi scientifiche solide e una validità universale (cfr. infatti il considerando 44). Ciò può, almeno in parte, dare conto delle incertezze nella loro sistemazione e spiegare come siano transitati durante la negoziazione dell'atto legislativo dal livello di basso rischio addirittura alle pratiche proibite. Qualora la tecnologia dovesse raggiungere una maturità che permette di validarne definitivamente i risultati, non si potrebbe che constatare la grande potenzialità invasiva delle capacità di interpretare gli stati interiori e di desumere l'attitudine emotiva dalle sembianze e dalle espressioni facciali. Queste considerazioni hanno portato a superare l'attenzione esclusivamente rivolta al caso degli applicativi usati in ambito commerciale, per decifrare le reazioni dei clienti e accordarvi le scelte in merito alla presentazione dei prodotti o alla disposizione della merce nel mondo reale, o per strutturare la grafica, i colori e gli slogan proposti nelle vendite online. A tale paradigma corrisponde ancora un obbligo di trasparenza, ossia di informare della presenza, altrimenti potenzialmente impercettibile, di questa rilevazione (art. 50, comma 3). Ma con una norma che coglie i rischi maggiori di certe tipologie di impiego è stato proibito di farne uso in ambito educativo e lavorativo (art. 5, comma 1, lett. f). Sono inoltre considerati ad alto rischio dispositivi come i poligrafi usati a scopi di polizia (Allegato III, n. 6, lett. b) o di controllo delle migrazioni (Allegato III, n. 7, lett. a).

Anche altri applicativi si presentano a geometria variabile, potendo essere collocati alternativamente tra i sistemi a basso rischio oppure, a seconda dell'impiego che ne viene fatto, costituire strumento per realizzare pratiche vietate rivolte alla manipolazione delle persone[8]. È il caso dei chatbot che si manifestano in molteplici ambiti e svolgono svariate funzioni: alcune del tutto innocue, come lo smistamento delle chiamate degli utenti da parte di siti e call center, cui si richiede solo di rendere nota la natura artificiale del risponditore; altre volte più delicate, specialmente quando si rivolgono ad adolescenti o altri soggetti vulnerabili e l'oggetto dell'interazione riguarda la sfera del benessere mentale, delle relazioni o dell'affettività[9].

In conclusione, tutti i sistemi che abbiano un'attitudine decettiva e integrino gli altri presupposti previsti dall'art. 5, lett. a) e b)[10], configurandosi come tecniche in grado di distorcere il comportamento e di pregiudicare la capacità di assumere decisioni informate, incontrano il divieto qui fissato. È tuttavia probabile che, nonostante l'ampliamento del perimetro di queste prescrizioni rispetto alla versione originaria, molte ipotesi restino al di sotto della soglia di rilevanza necessaria. Eventuali vuoti di protezione che dovessero situarsi tra le pratiche vietate, la disciplina dei sistemi ad alto rischio e quella dei sistemi appartenenti alla categoria sottostante, connotata da meri obblighi di trasparenza, potranno tuttavia essere compensati sfruttando tutte le possibili sinergie con altri corpi normativi.

3.1 … e complementarietà con altri ordini giuridici

I candidati ideali per comporre un quadro compiuto di tutele nel senso appena descritto sono costituiti dal diritto dei consumatori, dalla normativa per la protezione dei dati personali e dal cd. diritto delle piattaforme.

---

[7] Per un valore di mercato che era stato stimato in 90 miliardi di dollari entro il 2024: A. RENDA *et al.*, *Study to Support an Impact Assesment of Regulatory Requirements for Artificial Intelligence in Europe. Final Report*, 2021, 39 s.

[8] Le Commission Guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (AI Act), pp. 12 ss., chiariscono che l'accertamento del livello di rischio deve sempre farsi in concreto, sulla base delle funzionalità reali, potendo pertanto uno stesso sistema essere qualificato come vietato, ad alto rischio o a basso/medio rischio a seconda delle sue specifiche modalità di impiego.

[9] È il caso del chatbot Replika, interessato da un'indagine del Garante per la protezione dei dati personali (cfr. il provv. n. 39 del 2 febbraio 2023): attraverso la configurazione di un amico virtuale, permetteva conversazioni anche con soggetti minori di età e l'esposizione a contenuti giudicati inopportuni e pericolosi.

[10] Per un'analisi dettagliata della norma cfr. S. Orlando, *Regole di immissione sul mercato e "pratiche di intelligenza artificiale" vietate nella proposta di* Artificial Intelligence Act, in "Persona e mercato", 2022, 346 ss.

Nel primo plesso normativo si trovano sia norme puntuali sia discipline di portata più ampia, come quella sulle pratiche commerciali scorrette, che possono valere a contrastare attività algoritmiche abusive. Sotto il primo profilo, la modernizzazione del diritto dei consumi ha introdotto diversi obblighi di informazione per i professionisti, cui non corrisponde tuttavia, com'è noto, una capacità elevata di protezione dal rischio di manipolazione: sia per l'ampio ricorso che vi è stato fatto da parte del legislatore, al punto da generare un vero e proprio sovraccarico cognitivo, sia per una tipica reazione di indifferenza, empiricamente osservata nei destinatari, con un risultato controfinale rispetto alla *ratio* delle norme.

Per la comprensività dell'approccio, l'assetto entro cui potranno trovare sistemazione molte delle applicazioni algoritmiche che rimangono al di fuori dell'orizzonte dell'AI Act è piuttosto la disciplina delle pratiche commerciali scorrette. Essa appare l'ordine giuridico in grado di svolgere un'azione efficace di governo del mercato delle tecnologie, benché la sua ispirazione generale di tutela degli interessi economici dei consumatori renda meno adeguatamente presidiati gli altri diritti fondamentali messi a rischio dalle tecniche di manipolazione osservate.

Una casistica in formazione dimostra del resto come l'ampia configurazione delle nozioni di pratica commerciale e di decisione di consumo, articolate intorno all'offerta di servizi in cambio dell'estrazione dei dati dell'utente, abbia permesso di considerare illecita, e potenzialmente foriera di danni per la salute fisica e mentale, la funzione di personalizzazione del *feed* della piattaforma TikTok. La presentazione reiterata di video dove si esibiscono comportamenti autolesionistici in guisa di sfida, esponendovi in misura più accentuata proprio quegli utenti che risultano maggiormente "ingaggiati" dalla visione, è ritenuta idonea a falsare in misura apprezzabile il comportamento economico del consumatore, che si esprime anche attraverso lo scorrimento della sequenza di contenuti e la accresciuta permanenza sul social[11].

Ancora nell'ottica di valorizzare le interazioni proficue tra distinti plessi normativi, un riferimento essenziale è rappresentato dal GDPR. Il coinvolgimento di questo regolamento nella disciplina delle pratiche algoritmiche ad alto rischio dipende sia dall'ovvia constatazione che esse implicano molto spesso il trattamento di dati personali, e sono soggette perciò all'intero testo normativo; sia dalla presenza di previsioni specifiche riguardanti le decisioni automatizzate (artt. 13-15; art. 22). Non è un caso che prima di divenire oggetto di una disciplina apposita, nei termini del divieto o della sicurezza elevata, alcuni applicativi siano stati analizzati e giudicati nella loro liceità dal punto di vista del rispetto delle regole della *data protection*: con riferimento all'esistenza di una base normativa per il trattamento dei dati, alla validità del consenso, alla correttezza dell'informativa. Una rassegna, certamente non esaustiva, di queste ipotesi di doppia afferenza comprende i casi della piattaforma che offriva un servizio di elaborazione di profili reputazionali[12], dei progetti usati per scopi di polizia predittiva e di sorveglianza urbana[13], dei sistemi di riconoscimento facciale addestrati

[11] AGCM, provv. n. 31124 – TikTok cicatrice francese, 5 marzo 2024, in *Bollettino* n. 11 del 18 marzo 2024, pp. 67 ss.

[12] Si tratta del caso Mevaluate, che ha attraversato diversi gradi di giudizio: Garante per la protezione dei dati personali, provv. n. 488 del 24 novembre 2016; Trib. Roma, 4 aprile 2018, n. 5715; Cass., 25 maggio 2021, n. 14381, in "Diritto dell'informazione e dell'informatica", 2021, pp. 1001 ss., con nota di F. Bravo, Rating *reputazionale e trasparenza dell'algoritmo. Il caso "Mevaluate"*; Cass., 10 ottobre 2023, n. 28358, in "Nuova giur. civ. comm.", 2024, I, pp. 408 ss., con nota di N. Brutti, *Mito del consenso e* rating *reputazionale*. Sulla vicenda si v. anche, nell'ambito di un saggio dal più ampio respiro, G. Cerrina Feroni, *Intelligenza artificiale e sistemi di* scoring *sociale. Tra distopia e realtà*, in "Diritto dell'informazione e dell'informatica", 2023, pp. 11 ss.

[13] Garante per la protezione dei dai personali, provv. n. 5 dell'11 gennaio 2024 riguardante l'uso da parte del comune di Trento di sistemi di IA, sviluppati nell'ambito di alcuni progetti europei, che implicavano la raccolta di dati attraverso microfoni e videocamere di sorveglianza e la loro successiva elaborazione al fine di rilevare situazioni di pericolo per la pubblica sicurezza.

attraverso la raccolta indiscriminata di immagini pubblicate su siti e *account social* liberamente accessibili[14].

Quando, infine, i sistemi algoritmici siano integrati nelle piattaforme entra in gioco anche la disciplina del Digital Services Act; e al riguardo piuttosto incisivo sembra essere il divieto di ricorso ai *dark patterns*, ossia di strutturare le interfacce allo scopo di ingannare o manipolare i destinatari dei servizi (art. 25)[15].

Questi ragguagli sommari confermano come sia possibile osservare le applicazioni di IA da più angolazioni prospettiche, ciascuna in grado di offrire un contributo prezioso alla configurazione complessiva del loro statuto.

4. Normazione di sicurezza e standard tecnici

Il regolamento 2024/1689, lo si è notato a più riprese, costituisce una disciplina europea della sicurezza dei prodotti. A dispetto dell'enfasi posta sulla tutela dei diritti fondamentali e la difesa del carattere democratico della società[16], sin dalle enunciazioni fatte nei documenti ufficiali di avvio della Strategia europea, nei considerando iniziali e nella norma di apertura del testo ora approvato[17], tra i diversi modelli possibili di regolazione si è optato per questa soluzione tecnica, che affianca l'AI Act ai molti schemi già esistenti a livello europeo.

Ciò accade sia con la previsione di requisiti da osservare *ex ante*, orientati ai diversi tipi di rischio cui i sistemi autonomi possono esporre i diritti fondamentali, e alla vasta platea di regole procedurali sull'accertamento della conformità; sia nella scelta del meccanismo della coregolazione, che ai requisiti essenziali definiti nel testo legislativo affianca la predisposizione di standard tecnici diretti a specificarli.

Alcuni rilievi sono necessari con riguardo a entrambi i profili. Dal primo punto di vista, occorre osservare che i requisiti di sicurezza individuati sono i medesimi per entrambe le categorie di sistemi definite dall'art. 6, benché la discrepanza tra di esse, già evidenziata, sia qui destinata a riemergere. La prima è infatti quella più confacente al modello della *safety*, poiché gli interessi posti a rischio dai tipi che contiene sono, più verosimilmente, quelli della salute e dell'integrità fisica. La seconda classe, al contrario, è composta di applicazioni che più facilmente potranno interferire con diritti fondamentali o valori di natura diversa, come l'autonomia, la libertà, l'uguaglianza, la giustizia e il diritto di non essere discriminati. Integrare la tutela dei diritti fondamentali contro rischi diversi da quelli fisici nel modello della sicurezza dei prodotti, compendiato nella conformità e nell'apposizione del marchio CE, è un'operazione inedita e che presenta molteplici incognite[18].

---

[14] Si tratta dell'applicativo sviluppato dalla società Clearview, su cui il Garante per la protezione dei dai personali, così come numerose altre autorità europee, aveva avviato un'indagine conclusasi con il provv. n. 50 del 10 febbraio 2022.

[15] Sulle molteplici interazioni tra AI Act e DSA v. L. Ammannati e F. Di Porto, *L'intelligenza artificiale per la fornitura di servizi, di applicazioni e la produzione di regole:* Digital Services Act, Digital Markets Act e Artificial Intelligence Act, in *Intelligenza artificiale e diritto: una rivoluzione?*, cit., pp. 479 ss.

[16] Molto severo al riguardo è il giudizio di S. Pagliantini, *La base giuridica dell'AI Act ex art. 114 Tfue: l'intelligenza artificiale tra mercato e persona*, in "Foro it.", 2025, V, c. 22, che parla di "illusione ottica" e definisce la formula dell'IA antropocentrica come uno slogan "più proclamato che praticato".

[17] Per una puntuale rassegna dei passaggi chiave e della sequenza documentale improntati al personalismo e al rispetto dei valori fondativi dell'Unione europea cfr. A. Oddenino, *Intelligenza artificiale e tutela dei diritti fondamentali*, cit., pp. 171 ss.

[18] Cfr. le numerose critiche emerse fin dalla fase progettuale: M. Ebers, *Standardazing AI. The Case of the European Commission Proposal for an "Artificial Intelligence Act'*, in LA. DiMatteo, C. Poncibò e M. Cannarsa (a cura di), *The Cambridge Handbook of Artificial Intelligence. Global Perspectives on Law and Ethics*, Cambridge, CUP, 2022, pp. 321 ss.; C. Marenghi, *La proposta di regolamento UE sull'intelligenza artificiale e la regolazione privata: spunti critici in*

La regolazione fatta di requisiti tecnici da rispettare, inoltre, non solo è trasversale alle due categorie, ma è orizzontale in un senso più generale: si applica cioè ai sistemi di IA a prescindere dal loro settore di applicazione più specifico, che potrebbe essere quello della medicina, del mercato bancario, finanziario o assicurativo, del rapporto di lavoro, della gestione di funzioni pubbliche. Anche questa modalità è atipica rispetto alla legislazione in materia di sicurezza che è invece settoriale: e ciò genera l'interrogativo se il contesto di impiego possa essere indifferente alle scelte assunte nel momento in cui si progetta il sistema; ovvero se la destinazione d'uso del sistema non debba modellare in maniera più stringente la corrispondenza con le misure di sicurezza. Si tratta di un problema, invero, che non concerne solamente i cd. sistemi "general-purpose", per i quali si pone con maggiore acutezza; investe in realtà tutti i sistemi di IA che hanno un obiettivo determinato, ma potenzialmente specificabile in ragione dell'effettivo contesto di utilizzo.

L'ultimo punto, sempre a livello di analisi generalissima, riguarda le caratteristiche di dinamicità e di adattabilità lungo il ciclo di vita proprie dei sistemi basati sull'IA, che sono, del resto, gli elementi qualificanti della stessa definizione data dall'art. 3. Esse sono quasi del tutto ignorate dalla disciplina, se si eccettuano taluni obblighi di monitoraggio che permangono anche dopo la messa sul mercato, e la necessità di attivare un nuovo processo di conformità quando intervenga una modifica sostanziale (art. 43), che non riguarda però il caso dei cambiamenti dovuti all'apprendimento continuo quale modo operativo installato fin dall'inizio (lo precisa il comma 4)[19]. L'accento della regolamentazione riguarda infatti soprattutto la fase di progettazione ed è pertanto indirizzata principalmente al fornitore del sistema, l'equivalente del produttore di beni tangibili. La dimensione dell'utilizzatore che lo integra nella propria attività non è completamente assente, visto che alcuni obblighi insistono sul "*deployer*", secondo la nomenclatura adottata dall'AI Act (artt. 26 e 27). L'approccio di sicurezza, tuttavia, ancora una volta offusca la circostanza che si è di fronte a oggetti del tutto peculiari, più vicini alla matrice del servizio che a quello del prodotto, ai quali non si attaglia completamente un controllo in termini di conformità. Le qualità "strutturalmente mobili" delle applicazioni di IA sono insomma refrattarie e una gestione incentrata prevalentemente sul controllo *ex ante*[20].

In linea con la scelta per lo schema della sicurezza dei prodotti, viene poi accolto il metodo della co-regolazione, che nell'ambito del cd. Nuovo approccio richiede una collaborazione tra attori pubblici e attori privati: attraverso una delega agli organismi di standardizzazione, i requisiti generali fissati nell'AI Act saranno specificati da norme tecniche armonizzate, il cui rispetto, una volta pubblicate, darà luogo a una presunzione di conformità alla normativa (art. 40, comma 1). In sé apparentemente neutra, e largamente sperimentata in molti settori, questa forma di regolazione ibrida rischia di collidere con la natura affatto particolare dei sistemi di IA, che non sono invece prodotti come tutti gli altri.

Un primo problema attiene alla novità della tecnologia, cosicché la conoscenza scientifica sull'IA procede di pari passo con gli sviluppi normativi, rendendo complicato elaborare degli standard che presuppongono uno stato dell'arte più consolidato. Le competenze tecniche necessarie per regolare la progettazione e gli usi dell'IA e delineare strategie per i problemi di trasparenza, spiegabilità, giustizia algoritmica non sono ancora così compiute da poter essere consegnate a soluzioni operative.

---

tema di norme tecniche armonizzate, in "Diritto del commercio e degli scambi internazionali", 2021, pp. 575 ss.; M. Veale e F. Borgesius, *Demistifying the draft EU Artificial Intelligence Act*, in "Computer Law Review International", 4/2021, pp. 104 ss.; G. Resta, *Cosa c'è di 'europeo' nella proposta di Regolamento UE sull'intelligenza artificiale?*, in "Diritto dell'informazione e dell'informatica", 2022, pp. 340 ss.; M. Gornet e W. Maxwell, *The European approach to regulating AI through technical standards*, in "Internet Policy Review", 2024, vol. 13, pp. 1 ss.

[19] A. Oddenino, *Intelligenza artificiale e tutela dei diritti fondamentali*, cit., p. 186 s.

[20] A. Oddenino, *Intelligenza artificiale e tutela dei diritti fondamentali*, cit., p. 199.

Ma a creare perplessità è soprattutto il fatto di affidare una parte essenziale della regolazione a soggetti privati quali gli organismi di standardizzazione[21]. È noto come il processo di creazione delle norme tecniche non sia neutro, bensì intriso di scelte politiche tanto nel momento della costruzione quanto in quello successivo della diffusione[22]. Con la delega agli organismi di standardizzazione, si ripropongono dunque problemi già molto discussi nella letteratura in materia: circa un deficit di legittimazione democratica, poiché in questi organismi, che hanno lo statuto di organizzazioni di diritto privato, sono rappresentati soprattutto esponenti del mondo produttivo, nonostante vari tentativi di incrementare la partecipazione pubblica, della società civile e delle imprese medio-piccole; con riguardo alla scarsa trasparenza dei processi di decisione e al rischio di cattura regolatoria[23]; ancora, rispetto alla normazione dell'IA emerge un vero e proprio deficit di competenza, posto che al loro interno non c'è esperienza di confronto con i diritti fondamentali.

Se il modo di produzione normativa che combina atti legislativi e norme tecniche è tipico di molti comparti industriali, ciò non lo rende necessariamente adatto a occuparsi di scelte di valore e di bilanciamento tra interessi conflittuali, che saranno spesso la cifra caratteristica della regolazione in questa materia. Le decisioni assunte potranno riguardare, ad esempio, la definizione di soglie di performance accettabili dell'algoritmo, e dunque il livello atteso di precisione nel fare una diagnosi o nell'assistere una decisione di condanna in un procedimento penale. Oppure avranno a che fare con il modo di amministrare il problema della giustizia algoritmica, e se sia ammissibile che differenze statistiche, dotate di valore predittivo ma non direttamente pertinenti alla decisione da assumere o alla valutazione da compiere, siano riflesse nell'*output* dell'algoritmo. O, ancora, se si debba pretendere un'uguale distribuzione di risultati all'interno di gruppi aggregati intorno a caratteristiche protette, ancorché ciò comporti di rinunciare, almeno in parte, all'accuratezza del calcolo.

Questi dilemmi, invero, non vanno trattati come questioni puramente tecniche, bensì devono essere discussi nel circuito politico e democratico, dotati di risposte a livello legislativo ed, eventualmente, sindacati dai giudici qualora diventino oggetto di un reale contenzioso. Tanto che si fanno strada tentativi di resistere alla distribuzione di poteri e prerogative promossa dall'AI Act: con l'invito agli organismi di standardizzazione, incaricati di risolvere attraverso il tecnicismo nodi etici e giuridici, a limitarsi a indicare processi trasparenti con cui altri soggetti possano raggiungere le decisioni[24]; o a non servirsi degli standard per fissare soglie di accettabilità dei rischi e offrire regole di compromesso tra parametri in conflitto, ma soltanto per articolare diverse metodologie di misurazione e mitigazione, che i fornitori applicheranno sotto la loro responsabilità, e saranno infine sottoposte al vaglio delle autorità di sorveglianza e delle corti[25].
A queste criticità, già piuttosto pesanti, si aggiunge l'osservazione per cui la catena di controllo sarà nella maggior parte dei casi interamente affidata al settore privato, dallo sviluppo degli standard alla valutazione di conformità successiva[26]. Per la gran parte dei sistemi ad alto rischio, infatti, la conformità dei prodotti sarà asseverata attraverso meri controlli interni, che non coinvolgono organismi terzi indipendenti[27]. Anche qualora siano coinvolti gli enti notificati, mancherà comunque

---

[21] Sul punto G. Smorto, *Distribuzione del rischio e tutela dei diritti nel regolamento europeo sull'intelligenza artificiale*, in "Foro it.", 2024, V, cc. 214 ss.
[22] A. Solow-Niederman, *Can AI Standards Have Politics?*, in "UCLA Law Review", 2023, vol. 71, pp. 2 ss.
[23] Scenari preoccupanti emergono dall'indagine condotta dal Corporate Europe Observatory, di cui si riferisce in *Bias baked in. How Big Tech set its own AI standards*, 9.1.2025, https://corporateeurope.org/en/2025/01/bias-baked?utm_source=substack&utm_medium=email.
[24] J. Laux, S. Watcher e B. Mittelstadt, *Three pathways for standardisation and ethical disclosure by default under the European Union Artificial Intelligence Act*, in "Computer Law and Security Review", 2024 (53).
[25] M. Gornet e W. Maxwell, *The European approach to regulating AI through technical standards*, cit., p. 19.
[26] M. Gornet e W. Maxwell, *The European approach to regulating AI through technical standards*, cit., p. 7.
[27] In senso critico, U. Ruffolo e A. Amidei, *La regolazione* ex ante *dell'intelligenza artificiale tra gestione del rischio* by design*, strumenti di certificazione e "autodisciplina" di settore*, in *Intelligenza artificiale e diritto: una rivoluzione?*, cit., pp. 508 s.

un monitoraggio diretto da parte di agenzie pubbliche come avviene per altri prodotti considerati rischiosi quali le specialità farmaceutiche. E ciò aggrava il giudizio sul ruolo preponderante degli attori economici e tecnologici nella traiettoria di regolazione e nei processi di certificazione, rispetto allo spazio della discussione pubblica e del confronto politico e istituzionale.