Intelligenza artificiale e sovranità digitale: il modello europeo

Fernanda Faini*

ARTIFICIAL INTELLIGENCE AND DIGITAL SOVEREIGNTY: THE EUROPEAN MODEL
ABSTRACT: This document aims to analyse the relationship between European legislation and artificial intelligence, its distinctive features and the paradigms on which it is based. Specifically, this contribution intends to examine the philosophical-legal model of regulatory sovereignty that emerges in the light of European legislation on artificial intelligence. The European model of technological humanism shows an innovative approach that concerns, on the one hand, the relationship between law and technology and, on the other, the relationship between human beings and artificial intelligence. This model, based on legal protection by design, a risk-based approach, human oversight and explainability, marks an evolution in legal solutions and tools capable of innovating traditional paradigms.

KEYWORDS: artificial intelligence; digital sovereignty; European regulation; national regulation; technological humanism

ABSTRACT: Il saggio intende analizzare il rapporto tra regolazione europea e intelligenza artificiale, le caratteristiche distintive e i paradigmi su cui si basa. Nello specifico, il saggio esamina il modello filosofico-giuridico di sovranità regolatoria che emerge alla luce della legislazione europea in materia di intelligenza artificiale. Il modello europeo di umanesimo tecnologico mostra un approccio innovativo che riguarda, da un lato, il rapporto tra diritto e tecnologia e, dall'altro, il rapporto tra essere umano e intelligenza artificiale. Questo modello, basato sulla protezione giuridica fin dalla progettazione, su un approccio basato sul rischio, sulla supervisione umana e sulla spiegabilità, mostra un'evoluzione nelle soluzioni e negli strumenti giuridici in grado di innovare i paradigmi tradizionali.

PAROLE CHIAVE: intelligenza artificiale; sovranità digitale; regolazione europea; regolazione nazionale; umanesimo tecnologico

* *Professoressa associata di informatica giuridica (GIUR-17/A – Filosofia del diritto) presso il Dipartimento di Giurisprudenza dell'Università Telematica Pegaso. Mail: fernanda.faini@unipegaso.it.* *Il saggio è stato realizzato nell'ambito delle attività del Centro di Ricerca Interuniversitario "Centre for Law and Ethics of Innovation, Technology and Artificial Intelligence" (LEITAI), promosso dalle Università Telematica San Raffaele Roma, Università Telematica Pegaso e Università Telematica Universitas Mercatorum. Contributo sottoposto a doppio referaggio anonimo.*

protection by design e approccio basato sul rischio – 4.2. Il rapporto tra essere umano e tecnologia: supervisione umana e trasparenza algoritmica – 5. Conclusioni: riflessioni filosofico-giuridiche.

1. Intelligenza artificiale e diritto digitale

Il diritto è chiamato a governare la dimensione digitale, che connota in modo pervasivo l'esistenza contemporanea, al fine di tutelare i diritti e bilanciare interessi diversi e talvolta contrapposti. La dimensione attuale poggia su dati, algoritmi e tecnologie capaci di estrarne valore, come l'intelligenza artificiale, cui di conseguenza la scienza giuridica è tenuta a interfacciarsi.

Nello svolgere la sua funzione il diritto sconta le peculiarità dell'oggetto che mira a regolare, ossia l'intelligenza artificiale, costituita da beni intangibili, dati e algoritmi, diversi dalle *res corporales* cui il diritto è tradizionalmente abituato.

La società odierna, quale *data society*, è intimamente pervasa dai dati finendo per plasmare l'uomo stesso come un *data subject*[1]. Oggetti e soggetti diventano digitali; i beni si trasformano in servizi; il radicato paradigma della proprietà viene scalzato dal paradigma dell'accesso ai dati, ai servizi, alla propria esistenza digitale[2]. Di conseguenza diventa protagonista della contemporaneità l'intelligenza artificiale, la cui "anima" sono i dati e il "motore" gli algoritmi.

Pertanto, il diritto si trova a fare i conti con il peculiare oggetto di regolazione costituito dalla tecnologia, dominato da un ecosistema di regole diverso da quello giuridico, ossia le regole applicate dal codice informatico (*lex informatica*)[3], capaci di condizionare il comportamento dell'uomo, dal momento che rendono possibili o meno azioni e interazioni, definendone modi e vincoli, collegando effetti, determinando quali informazioni fornire all'utente e, di conseguenza, il grado di trasparenza e comprensibilità per chi le utilizza[4].

Nel determinare ciò che è possibile tecnologicamente, la *lex informatica* ha la capacità di condizionare ogni altra forma di regolazione, compresa quella giuridica, e l'uomo, per mezzo del diritto, deve essere capace di governarla, riuscendo a raggiungere un complesso equilibrio basato sulla flessibilità e sull'adattabilità, senza limitare le potenzialità dell'evoluzione tecnologica, ma, allo stesso tempo, fondato sulla prevedibilità e sulla certezza del diritto, senza determinare il dominio della tecnologia sulla regolazione[5].

Nel caso della tecnologia costituita dall'intelligenza artificiale alcune caratteristiche ontologiche incidono in modo particolarmente significativo nel rapporto con la regolazione giuridica.

Il regolamento UE 2024/1689 (*AI Act*) definisce un sistema di intelligenza artificiale (di seguito anche *Artificial Intelligence* o AI) come «*un sistema automatizzato progettato per funzionare con livelli di autonomia variabili e che può presentare adattabilità dopo la diffusione e che, per obiettivi espliciti o impliciti, deduce*

[1] In merito sia consentito il rinvio a F. FAINI, *Data society. Governo dei dati e tutela dei diritti nell'era digitale*, Milano, 2019.
[2] Cfr. L. FLORIDI, *La rivoluzione dell'informazione*, trad. it., Torino, 2012, 15.
[3] L. LESSIG, *The Law of the Horse: What Cyberlaw Might Teach*, in *Harvard Law Review*, 113, 1999, 501-546.
[4] L'incisiva locuzione "code is law", usata da Lessig, evidenzia proprio l'aspetto regolatorio insito nel codice informatico nella dimensione digitale; cfr. L. LESSIG, *Code and Other Law of Cyberspace*, New York, 1999.
[5] Cfr. V. FROSINI, *Il diritto nella società tecnologica*, Milano, 1981.

dall'input che riceve come generare output quali previsioni, contenuti, raccomandazioni o decisioni che possono influenzare ambienti fisici o virtuali»[6].

L'autonomia, infatti, connota e differenzia gli algoritmi di intelligenza artificiale dalla logica degli algoritmi tradizionali, dove l'impostazione deterministica *if this*, *than that* indica un funzionamento del tutto predeterminato nel programma, che prevede gli *input* e gli *output* e si limita ad applicare regole informatiche predefinite ed espresse in linguaggio di programmazione. Nell'apprendimento automatico (*machine learning* e *deep learning*)[7] l'uomo fornisce alla macchina un metodo di apprendimento da applicare ai dati cui ha accesso, per estrarre automaticamente le indicazioni e le nozioni necessarie per l'assunzione di una determinazione, analizzando grandi quantità di dati ed apprendendo i parametri numerici, necessari per adottare decisioni e utilizzarli nella successiva fase di esecuzione. La rappresentazione matematico-numerica, ossia il modello generato nella fase di apprendimento o *training*, solitamente non è direttamente intelligibile da parte dell'essere umano (*black box*), aspetto che rileva particolarmente sotto la lente giuridica[8]. Proprio in questo aspetto si scorge la natura "intelligente" della macchina, che è capace di apprendere attraverso i dati in modo autonomo.

L'intelligenza artificiale si affida a inferenze e connessioni tra dati e poggia sull'approccio statistico e probabilistico, determinando talvolta difficoltà di comprensione circa le motivazioni (la *ratio*, il "perché") delle risposte fornite[9]. Pertanto garantire trasparenza algoritmica può essere particolarmente complesso a fronte di una congenita opacità degli algoritmi, che si declina in un'opacità strutturale, derivante dal funzionamento degli stessi e dal fatto che resta non comprensibile persino ai programmatori l'iter logico seguito dalla macchina per giungere al risultato partendo dai dati a disposizione, cui si somma l'opacità linguistica, dovuta al linguaggio informatico e non a quello naturale[10].

Di conseguenza emergono problematiche giuridiche strettamente correlate alle caratteristiche ontologiche che connotano l'intelligenza artificiale, quali il meccanismo di inferenze e correlazioni su cui si basano gli algoritmi, la correlata difficile intelligibilità da parte dell'essere umano (*black box*), la connessa difficoltà di motivazione dei risultati cui perviene la macchina, oltre a possibili errori e *bias*, forieri di potenziali discriminazioni, disuguaglianze, disparità[11]. La tecnologia, infatti, si staglia quale protagonista solo teoricamente "neutrale", dal momento che nella sua valenza strumentale acquisisce il significato che gli uomini le conferiscono nell'utilizzarla; le tecnologie sono "lame a doppio taglio" usando l'efficace configurazione di Bauman[12] e, di conseguenza, nel contesto digitale possono emergere le asimmetrie e le disparità tipiche della società analogica.

Nella regolazione della tecnologia il diritto si trova di fronte anche a un mutamento della dimensione relativa allo spazio, che subisce fortemente l'impatto della rivoluzione digitale; la dimensione globale di riferimento, priva di confini territoriali e "sovrani", determina la rilevanza in materia del diritto sovranazionale. La *ratio* di tale aspetto si rinviene nell'esigenza di un approccio olistico per affrontare la

[6] Art. 2, par. 1, n. 1, reg. UE 2024/1689.

[7] Il deep learning è basato su reti neurali artificiali a molti strati; cfr. F. LAGIOIA, G. SARTOR, *L'intelligenza artificiale per i diritti dei cittadini: il progetto Claudette*, in *Ragion pratica*, 1, 2020, 88 ss.

[8] Infra, § 4.2.

[9] Cfr. P. MORO, *Algoritmi e pensiero giuridico. Antinomie e interazioni*, in *Rivista di diritto dei media*, 3, 2019, 19.

[10] Infra, § 4.2.

[11] G. CORASANITI, *Tecnologie intelligenti. Rischi e regole*, Milano, 2023, 63 ss.

[12] Z. BAUMAN, D. LYON, *Sesto potere. La sorveglianza nella modernità liquida*, trad. it., Roma-Bari, 2015, 86 ss.

dimensione digitale, dal momento che la realtà globale della rete implica un mutamento nei confini geografici della regolazione e rende necessario arrivare a soluzioni condivise in merito alla protezione dei diritti; la veste sovranazionale è necessaria a garantire efficacia alle norme, evitando la tensione altrimenti fisiologica tra la connotazione globale delle questioni e la dimensione territoriale delle disposizioni da applicare.

Inoltre lo spazio/non spazio digitale è capace di ridefinire sfera pubblica e privata e dare vita a nuove geometrie di potere, in cui gli Stati sono depotenziati ed erosi nel loro potere sul mondo digitale a causa del territorio globale di azione e della presenza delle aziende *Big Tech*, "controllori del pedaggio di accesso alla vita digitale", favoriti dalla stessa dimensione sovranazionale, dalla capacità di utilizzare le tecnologie, dalla tutela della libertà di impresa e dello sviluppo economico, dalle maglie ampie o dall'assenza di disposizioni adeguate al mutato contesto. Per mezzo degli strumenti dell'autonomia contrattuale privata, le aziende *Big Tech* hanno avuto la capacità di dominare la dimensione globale, regolando l'accesso ai servizi e alle utilità della rete, incidendo così sui diritti e sulle libertà dei singoli.

I colossi tecnologici concretamente producono regole in un contesto di disintermediazione, in cui con una provocazione danno vita un ordine algoritmico in cui si atteggiano da potere legislativo (dettano le regole), esecutivo (le applicano) e giudiziario (giudicano le violazioni con sanzioni quali la sospensione o la disattivazione dell'*account*). Si tratta di regole ampiamente accettate, seppur più o meno consapevolmente, capaci di influire sulla vita degli individui quanto le norme con efficacia vincolante prodotte dagli Stati; i nuovi "territori" costituiti dalle piattaforme digitali sono percepiti dalla collettività come spazi giuridici pubblici, sebbene privi di legittimazione democratica e guidati non dall'interesse pubblico, ma dal profitto economico[13].

Alla luce di tale contesto di riferimento, il contributo intende analizzare il rapporto tra regolazione europea e intelligenza artificiale, le sue peculiarità e i paradigmi su cui si erge. In particolare, il contributo è teso ad esaminare le caratteristiche peculiari del modello filosofico-giuridico di sovranità regolatoria, che emerge alla luce della regolazione europea in materia di intelligenza artificiale. Il modello europeo di umanesimo tecnologico mostra un approccio innovativo che riguarda, da un lato, il rapporto tra diritto e tecnologia e, dall'altro, il rapporto tra esseri umani e intelligenza artificiale, mostrando un'evoluzione delle soluzioni e degli strumenti giuridici in grado di innovare i paradigmi tradizionali.

2. Sovranità regolatoria europea

Nella società digitale contemporanea, in cui il diritto è chiamato a confrontarsi con le peculiarità dell'oggetto di regolazione e con mutate geometrie di potere, negli ultimi anni prende forma un articolato *framework* giuridico a livello europeo diretto a disciplinare la dimensione digitale della vita umana, formato inizialmente prevalentemente da atti di *soft law*[14] e in seguito da un insieme di regolamenti quali il *Cybersecurity Act* (regolamento UE 2019/881), il *Data Governance Act* (regolamento UE 2022/868), il *Digital Markets Act* (regolamento UE 2022/1925), il *Digital Services Act* (regolamento UE 2022/2065), il *Data Act*

[13] In merito all'inquadramento delle piattaforme digitali quali spazi giuridici pubblici si rinvia alla vasta letteratura in materia; cfr., inter alia, H. MICKLITZ, O. POLLICINO, A. REICHMAN, A. SIMONCINI, G. SARTOR, G. DE GREGORIO (a cura di), *Constitutional Challenges in the Algorithmic Society*, Cambridge, 2021.

[14] Infra, nota 26.

(regolamento UE 2023/2854), il *Cyber Resilience Act* (regolamento UE 2024/2847), il *Cyber Solidarity Act* (regolamento UE 2025/38) e l'*Artificial Intelligence Act* (regolamento UE 2024/1689; di seguito anche *AI Act*).

Nella propria produzione normativa l'Unione europea, oltre ad esercitare un ruolo guida al fine di garantire un approccio comune degli Stati membri e realizzare un ecosistema digitale sicuro, competitivo, inclusivo e antropocentrico, mostra la volontà di costruire un modello filosofico-giuridico di governo della tecnologia con cui affermare una sovranità digitale regolatoria rispetto ai poteri privati e ad altri blocchi geopolitici.

La società algoritmica, infatti, pone il problema della sovranità digitale, legata alla capacità dello Stato di avere accesso, controllo, gestione e protezione di dati, software e infrastrutture informatiche all'interno del proprio contesto geopolitico, senza dipendere da potenze, piattaforme o aziende esterne; la sovranità statale passa dalla sovranità digitale, fondamento della moderna sovranità. La sovranità digitale si esercita sicuramente sui dati, sulle tecnologie, sulle infrastrutture, sul controllo relativo alle attività compiute nella dimensione digitale, ma anche sulle norme e sulla connessa capacità di incidere con la regolazione. Proprio nelle diverse componenti su cui può esercitarsi si stagliano gli attuali volti della sovranità digitale a livello geopolitico internazionale con un diverso approccio tra Stati Uniti, Cina e, appunto, Unione europea.

Il modello statunitense dà vita a una sovranità delle piattaforme private, dotate di un ruolo globale e del dominio economico sul mercato, caratterizzata da un limitato intervento normativo a livello unitario e dalla cooperazione pubblico/privato, che si declina nella collaborazione dello Stato con aziende e poteri privati[15].

Il modello cinese, invece, costruisce una sovranità statale che si erge su un controllo diretto, totale e pervasivo dello Stato su dati, infrastrutture e contenuti e si articola nell'approccio centralizzato, nel controllo e nella censura in Internet (*firewall*, filtraggio, tracciamento) e nella costruzione di piattaforme interne autoctone, accompagnata parallelamente dal blocco delle grandi piattaforme digitali esterne. In tal modo il sistema cinese mira all'autonomia tecnologica, riducendo la dipendenza dall'esterno, oltre a ricercare sicurezza e protezione dell'ordine sociale e politico che si realizza anche con la sorveglianza di massa sugli utenti (*smart cities*, *social scoring*, riconoscimento facciale)[16].

Tra approccio liberista statunitense e logica centralizzata cinese, l'Unione europea abbraccia una terza via e costruisce una sovranità regolatoria, cercando di esprimere capacità di regolamentazione, ponendo condizioni giuridiche per mezzo dell'insieme delle norme dedicate alla dimensione digitale; l'Unione europea cerca così di stabilire le regole del gioco a livello globale. Il modello europeo fa leva sulla tutela dei diritti, del diritto e della democrazia rispetto al potere delle *Big Tech*, esprimendo un'autonomia strategica e mirando a realizzare un mercato digitale equo e sicuro, capace di dare vita a un modello antropocentrico che possa conquistare la fiducia degli utenti, tutelati nei propri diritti fondamentali[17]. Tale approccio sil

[15] Al riguardo si segnalano gli standard tecnici e le policy non vincolanti statunitensi come il NIST AI Risk Management Framework (2023), standard volontario e non prescrittivo, adattabile a diversi settori e dimensioni aziendali.

[16] In relazione al modello cinese si segnalano le Interim Measures for the Management of Generative AI Services (2023), secondo cui i contenuti generati dall'AI devono essere conformi ai "valori socialisti fondamentali" ed è vietato qualsiasi output che possa minare il potere statale o la stabilità sociale.

[17] A. D'ATTORRE, *La sovranità digitale. Poteri privati, intervento pubblico e diritti individuali nel cyberspazio*, in T. CASADEI, S. PIETROPAOLI (a cura di), *Diritto e tecnologie informatiche*, 2° ed., Trento, 2024, 313-324.

lega al cosiddetto *effetto Bruxelles*, ossia alla capacità dell'Unione Europea di regolare i mercati globali in modo unilaterale grazie a un adeguamento volontario delle aziende alle regole europee al fine di evitare costi[18]. Tale capacità di influenza è in realtà ancora da verificare nel caso dei regolamenti più recenti come quello sull'intelligenza artificiale, che sconta anzi l'ombra del *chilling effect*, ossia il fatto che le norme europee possano scoraggiare le imprese dall'intraprendere attività innovative nel mercato unico per timore di sanzioni significative e costi di *compliance*.

In relazione alla via scelta dall'Unione europea e alla volontà di orientare altri blocchi geopolitici, rileva la Convenzione quadro sull'AI del Consiglio d'Europa (trattato di Vilnius, 2024), primo trattato internazionale giuridicamente vincolante in materia, che a differenza dell'AI Act ha una portata globale, essendo aperta anche a stati extra-europei.

La sovranità regolatoria emerge da una serie di aspetti che sinergicamente danno vita al modello giuridico europeo, teso a realizzare un diritto sostenibile nella dimensione digitale e, a tal fine, capace di raggiungere un saggio bilanciamento tra diritti e interessi diversi in conformità ai valori etici e ai principi giuridici che caratterizzano la tradizione europea.

In primo luogo, la finalità di porre le regole del gioco a livello mondiale esercitando la propria sovranità regolatoria emerge dalla scelta dell'Unione europea di avvalersi in modo crescente di atti normativi dotati di peculiare *vis* normativa come i regolamenti, ponendosi come pioniera in ambiti come l'intelligenza artificiale.

La sovranità regolatoria europea emerge altresì nella portata extraterritoriale delle norme che ricalca l'effetto trasformativo del regolamento europeo 2016/679 *General Data Protection Regulation* (il richiamato *Brussels Effect*). L'*Artificial Intelligence Act* (regolamento UE 2024/1689) mostra, infatti, un ambito di applicazione soggettivo particolarmente esteso, che, sulle tracce del precedente regolamento 2016/679 sulla protezione dei dati personali, esercita un effetto extraterritoriale rispetto ai confini europei, dal momento che si applica ai fornitori che immettono sul mercato o mettono in servizio sistemi di AI o modelli di AI per finalità generali nell'Unione, indipendentemente dal fatto che siano stabiliti o ubicati nell'Unione o in un paese terzo, e ai fornitori e *deployer* di sistemi di AI che hanno il loro luogo di stabilimento o sono situati in un paese terzo, laddove l'*output* prodotto dal sistema di AI sia utilizzato nell'Unione[19].

Al riguardo preme precisare che la proiezione globale dell'AI Act solleva complessi profili di conflitto di leggi. L'imposizione unilaterale di regole europee su attori globali potrebbe scontrarsi con giurisdizioni (come quella statunitense o cinese) che adottano criteri di sicurezza o di governance divergenti. In assenza di un coordinamento internazionale che la *Framework Convention* del Consiglio d'Europa mira in parte a colmare, si delinea una competizione tra sovranità diverse, in cui l'Europa tenta di imporre la propria regolazione come standard che possa guidare la democrazia digitale.

Altra leva della sovranità regolatoria europea è l'*enforcement* costruito dall'AI Act, che poggia su un solido meccanismo sanzionatorio basato sul *turnover* annuo globale, con sanzioni amministrative pecuniarie che possono raggiungere i 35 milioni di euro o il 7% del fatturato mondiale totale annuo dell'esercizio precedente, a seconda di quale sia l'importo più elevato per le violazioni relative alle pratiche vietate[20]. Questo

[18] Cfr. A. BRADFORD, *The Brussels Effect: How the European Union Rules the World*, in *Oxford University Press*, Oxford, 2020.
[19] Art. 2 reg. UE 2024/1689.
[20] Art. 99 reg. UE 2024/1689.

approccio assicura che il costo della non-conformità superi i potenziali benefici economici, agendo da potente deterrente anche per le Big Tech globali.

L'operatività tecnica del sistema è delegata al binomio tra standard armonizzati e organismi notificati, che traduce il piano normativo in requisiti tecnici applicabili. La Commissione Europea ha conferito mandato agli organismi europei di normazione di elaborare standard tecnici dettagliati su pilastri critici quali la robustezza, l'accuratezza e la cybersicurezza. L'adozione di tali standard consente ai fornitori di beneficiare di una presunzione di conformità, semplificando l'iter di immissione sul mercato[21]. Tuttavia, per i sistemi classificati ad alto rischio, la valutazione della conformità non è lasciata all'autocertificazione, ma è demandata all'intervento degli organismi notificati: enti terzi accreditati che agiscono come garanti e "filtri di qualità" *ex ante*, assicurando l'allineamento tecnologico ai valori giuridici dell'Unione prima della commercializzazione.

La sovranità regolatoria si manifesta anche nella costruzione di una solida *governance* tracciata dai regolamenti europei, che prevedono l'istituzione di Comitati e gruppi di esperti a livello sovranazionale, dedicati a specifici ambiti di azione o tecnologie (intelligenza artificiale, dati, servizi e mercati digitali) e chiamati a fornire assistenza e consulenza alla Commissione europea nell'applicazione uniforme dei regolamenti, quali l'*European Artificial Intelligence Board*[22], l'*European Data Innovation Board*[23], l'*European Board for Digital Services*[24], l'*high-level group for the Digital Markets Act*[25]. Nell'applicazione concreta ed omogenea delle previsioni giuridiche e nel bilanciamento tra diritti e interessi diversi, il ruolo di questi organismi è fondamentale al fine di rispondere alle istanze di flessibilità e adattabilità, necessarie affinché le regole giuridiche possano essere efficaci, siano capaci di affrontare l'evoluzione tecnologica e rispondere all'esigenza di creare un diritto sostenibile.

La costruzione di una g*overnance* dedicata alla dimensione digitale, articolata nei diversi organismi sovranazionali, evidenzia la volontà di porre le regole del gioco recuperando uno spazio giuridico pubblico a protezione dei diritti della persona, perso a causa del potere assunto concretamente dai soggetti privati.

Non mancano le criticità di questa impostazione; il rischio, infatti, consiste nel pericolo di far pendere la bilancia verso l'uno o l'altro diritto a protezione del quale o verso la tecnologia per la quale lo specifico organismo è istituito e conseguentemente si muove, perdendo l'approccio olistico necessario per affrontare la complessità della realtà digitale, caratterizzata dalla presenza di interessi diversi e dalla necessità di trovare un complesso equilibrio tra gli stessi.

Pertanto la previsione di organi sovranazionali chiamati a contribuire all'applicazione efficace ed omogenea delle norme determina la necessità conseguente di una sinergica cooperazione tra loro, superando l'ambito settoriale di azione che rischia altrimenti di far smarrire il richiamato approccio olistico necessario alla luce delle evidenti intersezioni tra i diversi regolamenti.

La sovranità regolatoria può essere esercitata solo attraverso la costruzione di un diritto sostenibile, esigenza che si manifesta nell'eterogeneità degli atti che ospitano le regole, che si coniuga all'esaminata eterogenea *governance* in materia: non solo *hard law*, ma anche regole di *soft law* prodotte da soggetti

[21] Art. 40 reg. UE 2024/1689.

[22] Art. 65 ss., titolo VI, reg. UE 2024/1689, che nella governance relativa all'AI accanto al Board prevede l'European AI Office, l'Advisory forum e un Comitato scientifico composto da esperti indipendenti.

[23] Art. 29 ss., reg. UE 2022/868 «Data Governance Act».

[24] Art. 61 ss., reg. UE 2022/2065 «Digital Services Act».

[25] Art. 40, reg. UE 2022/1925 «Digital Markets Act».

istituzionali, autorità indipendenti, organismi, comitati, gruppi di esperti nominati *ad hoc*; la complessità dei fenomeni è affrontata con strumenti eterogenei prodotti non solo dalle istituzioni pubbliche competenti, ma anche da soggetti privati[26].

Come precedentemente esaminato[27], al fine di governare la tecnologia, infatti, il diritto deve tener conto ed integrare ecosistemi di regole diverse, come quelle informatiche, che dettano "legge" nella dimensione digitale, e, altresì, quelle poste nell'ambito dell'autonomia privata dai colossi tecnologici, che regolano i "territori" delle estese piattaforme digitali su cui esercitano il proprio dominio[28]. Di conseguenza, la regolazione necessita di una genesi *multilevel*, di un approccio *multistakeholder* e dell'integrazione di fonti eterogenee, al fine di avvalersi di una corresponsabilità da parte dei differenti produttori di regole, giuridiche o meno, della nostra contemporaneità. Pertanto, accanto alle norme poste dalle autorità pubbliche, rilevano forme di *co-regulation* e di *self-regulation*, dove le regole sono poste da parte degli stessi destinatari delle stesse, come le aziende *Big Tech* che danno vita a una sorta di contemporanea *lex mercatoria*; gli Stati sono chiamati a nuove forme di cooperazione e collaborazione.

L'eterogeneità della tipologia degli atti si motiva proprio alla luce delle caratteristiche del peculiare oggetto di regolazione: la tecnologia è connotata dalla specificità tecnica, dal progresso scientifico e dall'evoluzione costante, che esigono qualificate competenze e comportano l'esigenza di atti puntuali, regole tecniche e strumenti flessibili, mentre la norma strettamente intesa si caratterizza per essere generale e astratta, frutto di un lungo iter scaturente dal processo democratico e tesa a durare nel tempo. Di conseguenza, è necessario un complesso articolato di regole diverse, atte a costruire una disciplina complessivamente sostenibile (e, dunque, efficace) e, a tal fine, tese a garantire flessibilità e adattabilità allo sviluppo e ai cambiamenti della tecnologia, ma capaci altresì di assicurare prevedibilità e certezza, principi solidi del diritto.

Nel quadro europeo in materia, in cui si esprime la sovranità regolatoria, emerge questa tensione continua fra flessibilità e prevedibilità, al fine di garantire efficacia ma anche certezza, tutelando la persona e riuscendo a trovare un virtuoso equilibrio tra interessi diversi.

L'esigenza di flessibilità e adattabilità rispetto all'incessante sviluppo tecnologico e la conseguente necessità di aggiornamento determinano meccanismi quali nell'*Artificial Intelligence Act* l'obbligo generale di revisione del regolamento e la procedura di verifica e modifica degli allegati connessi all'individuazione delle categorie dei sistemi ad alto rischio, anche al di fuori del procedimento legislativo ordinario altrimenti necessario; nella stessa ottica è possibile leggere le previsioni relative alle *regulatory sandboxes,* utili ad adattare le disposizioni a una realtà particolarmente complessa[29], e il ricorso allo strumento dei

[26] Con il termine hard law si intende un sistema di regole giuridiche vincolanti, dotate di piena obbligatorietà in senso giuridico, suscettibili in caso di violazione di applicazione giudiziaria, mentre con soft law un sistema di regole non precettive e non vincolanti, comunque caratterizzato da un diverso grado di persuasività, ossia in grado di svolgere effettivamente una funzione di orientamento e di indirizzo nei confronti dei destinatari, sebbene non suscettibili di attuazione giudiziaria.

[27] Supra, § 1.

[28] A. SIMONCINI, *Verso la regolamentazione della Intelligenza Artificiale. Dimensioni e governo*, in *BioLaw Journal*, 2, 2021, 412.

[29] L'attivazione delle regulatory sandboxes consente agli Stati membri, per un periodo di tempo limitato e sotto il controllo delle autorità competenti nazionali, di sviluppare e sperimentare sistemi di IA innovativi, ai fini di una successiva immissione nel mercato; art. 57 ss., reg. UE 2024/1689.

codici di buone pratiche e codici di condotta, atti a contribuire alla corretta applicazione delle norme[30]. Al riguardo, in relazione a modelli di intelligenza artificiale molto diffusi, quali quelli di AI generativa (es. ChatGPT), significativi esempi di tale esigenza di flessibilità sono il *General-Purpose AI Code of Practice*, documento volontario formalmente non vincolante, pubblicato il 10 luglio 2025 e redatto da esperti indipendenti con input da oltre 1.000 stakeholder come strumento di autoregolazione, e le successive *Guidelines on the Scope of Obligations of Providers of General-Purpose AI Models under the AI Act*, pubblicate dalla Commissione europea il 18 luglio 2025 per assistere i fornitori di modelli di AI per finalità generali ad adempiere agli obblighi del regolamento europeo.

L'esigenza di prevedibilità e certezza del diritto, insieme alla necessità di garantire omogeneità di soluzioni, determinano invece il cambiamento stesso di approccio da parte dell'Unione europea con la tendenza a "inasprire" la forza degli atti normativi assunti (da *soft law* ad *hard law*), al fine di garantirne effettività; è il caso del regolamento europeo che con la sua maggiore *vis* sugli Stati membri prende prevalentemente il posto della direttiva nel *framework* giuridico dedicato alla dimensione digitale[31] e, altresì, la previsione di sanzioni proporzionate al fatturato annuo globale, al fine di realizzare un apparato sanzionatorio effettivo, proporzionato e dissuasivo, che permetta di rispondere in modo certo in caso di violazioni[32].

Al riguardo occorre precisare che anche se viene fatta la scelta della tipologia di atto regolamentare non si rinuncia comunque all'eterogeneità di atti, al fine di garantire effettività alle regole poste: in tali disposizioni, infatti, si rinvia anche ad atti di *soft law* necessari per l'applicazione, come i richiamati codici di condotta. Inoltre, seppur lo strumento formale scelto per regolare la dimensione digitale sia il regolamento, oltre agli esaminati meccanismi di flessibilità, sono lasciati ampi margini di applicazione agli Stati membri, che equilibrano la maggior forza dell'atto con l'esigenza di garantire anche adattabilità ed efficacia, seppur tale profilo possa causare anche incertezze e potenziali difformità nell'attuazione, finendo per andare in direzione opposta rispetto a quella uniformazione necessaria ad esprimere sovranità regolatoria europea in materia[33].

3. Sovranità regolatoria nazionale

Il modello europeo, teso a realizzare un diritto sostenibile nella dimensione digitale e per quanto attiene all'intelligenza artificiale in particolare, mostra pertanto la volontà di esprimere una sovranità regolatoria in materia, differenziandosi dal modello statunitense e cinese.

E, seppur la dimensione sovranazionale sia quella adeguata alla regolazione in materia, non mancano atti a livello nazionale, per lo più strategici e stimolati dalla stessa Unione europea, cui da ultimo si è però affiancato un vero e proprio atto legislativo, cui occorre dedicare qualche riflessione.

[30] Art. 56 ss. e art. 95 ss., reg. UE 2024/1689.

[31] Tale tendenza ad abbandonare la direttiva in favore del regolamento era già in atto, si pensi ai regolamenti europei 2016/679 sui dati personali e 2018/1807 sui dati non personali.

[32] Art. 99 ss., reg. UE 2024/1689.

[33] Nell'AI Act, ad esempio, l'individuazione dei sistemi a rischio inaccettabile si connota per la presenza di concetti indeterminati ed interpretabili, che implica flessibilità applicativa e margini di manovra a favore degli Stati, ma anche possibile incertezza e correlate potenziali difformità nell'attuazione.

Dopo una serie di strategie dedicate, la prima del 2018 con il Libro Bianco "*L'intelligenza artificiale al servizio del cittadino*" per poi arrivare alla Strategia italiana del 2020 e quella attuale del 2024-2026[34], è stata approvata una legge nazionale in materia di intelligenza artificiale, la legge 132/2025, che interviene in diversi ambiti.

Oltre a definizioni, ambito di applicazione, finalità e principi[35], le norme spaziano su diversi aspetti: si prevedono disposizioni di settore in ambiti significativi (ambito sanitario; lavoro; professioni intellettuali; pubblica amministrazione; attività giudiziaria)[36]; è prevista la necessità di una strategia nazionale[37]; si designano quali autorità nazionali per l'intelligenza artificiale con correlata divisione di competenze AgID, competente per promuovere l'innovazione e lo sviluppo dell'AI, e l'Agenzia per la cybersicurezza nazionale (ACN), competente per la vigilanza e la promozione e sviluppo dell'AI relativamente alla cybersicurezza[38]; si prevedono una serie di ampie deleghe al Governo[39]; sono dedicate disposizioni alla tutela del diritto di autore[40]; si introducono sanzioni penali, come l'illecita diffusione di contenuti generati o alterati con sistemi di intelligenza artificiale[41].

Questa operazione mostra, a somiglianza di quanto avviene a livello europeo, la volontà di esprimere una sovranità regolatoria nazionale, pur dovendo necessariamente rimanere nei limiti di quanto previsto a livello europeo dall'*AI Act*, che in quanto regolamento si applica direttamente agli Stati membri, disapplicando eventuali norme nazionali in contrasto.

In linea con l'approccio europeo e con l'*AI Act*, la legge nazionale intende promuovere un utilizzo corretto, trasparente e responsabile, in una dimensione antropocentrica, dell'intelligenza artificiale, volto a coglierne le opportunità e garantire la vigilanza sui rischi economici e sociali e sull'impatto sui diritti fondamentali[42]; l'atto normativo precisa in modo opportuno che le norme si interpretano e si applicano conformemente al regolamento (UE) 2024/1689: del resto, non potrebbe essere altrimenti data la tipologia di atti normativi di cui si parla[43].

Ma, nonostante ciò, non mancano criticità, che meritano qualche riflessione proprio in relazione alla sovranità regolatoria.

Al riguardo, infatti, la Commissione europea aveva trasmesso all'Italia il 5 novembre 2024 un parere circostanziato (C(2024) 7814) circa potenziali criticità e sovrapposizioni tra il disegno di legge nazionale e il

[34] Si tratta dei seguenti atti: il Libro Bianco "L'intelligenza artificiale al servizio del cittadino", curato dall'AgID e dalla Task Force sull'IA composta da esperti nel marzo 2018; la Strategia italiana per l'intelligenza artificiale 2020, basata sulle proposte elaborate dal gruppo di esperti sull'AI nominato dal Ministero dello Sviluppo Economico 2019; la Strategia italiana per l'intelligenza artificiale 2024-2026, prodotta da parte del Comitato di Coordinamento per l'aggiornamento delle strategie sull'utilizzo dell'AI, istituito presso il Dipartimento per la Trasformazione digitale, composto da 13 componenti nominati ad ottobre 2023.

[35] Capo I, legge 132/2025.

[36] Capo II, legge 132/2025.

[37] Art. 19, legge 132/2025.

[38] Art. 20, legge 132/2025.

[39] Art. 24, legge 132/2025.

[40] Art. 25, legge 132/2025.

[41] Art. 26, legge 132/2025.

[42] Art. 1, comma 1, legge 132/2025.

[43] Art. 1, comma 2, legge 132/2025.

regolamento europeo *AI Act*, solo in parte superate nelle modifiche apportate lungo l'iter parlamentare[44]. Inoltre, perplessità emergono sull'opportunità stessa di una legge nazionale in un momento in cui il regolamento non è ancora pienamente applicato e con cui possono aprirsi pericolose discrasie; quanto meno la tempistica non pare congrua. Più ampiamente, alla luce del modello europeo di sovranità regolatoria che l'Unione europea, come esaminato, cerca di affermare, la stessa approvazione di una norma nazionale può risultare critica, perché può portare a una potenziale differenziazione con altri Stati e quindi a una frammentazione, in direzione diametralmente opposta alle finalità europee di uniformazione, che vengono perseguite con una *vis* normativa quale quella del regolamento, dotato anche di effetti extraterritoriali proprio al fine di imporsi a livello globale.

Le perplessità aumentano alla luce delle ampie deleghe al Governo e della clausola di invarianza finanziaria, secondo cui dall'attuazione non devono derivare nuovi o maggiori oneri a carico della finanza pubblica, che rischia di rendere oltretutto le disposizioni mere dichiarazioni di intenti, senza che si traducano in norme effettive ed omogeneamente attuate.

Anche la prevista *governance* duale con la designazione di AgID e ACN come autorità nazionali non risulta convincente, alla luce di una prevedibile difficoltà di distinzione delle rispettive competenze e conseguenti conflitti e sovrapposizioni. Lo snodo della *governance* è però cruciale per costruire la sovranità regolatoria europea. Peraltro, date le rilevanti competenze già agite da tali autorità e la complessità dell'intelligenza artificiale, sommare questa competenza ad organismi esistenti potrebbe non rivelarsi la soluzione ottimale al momento dell'implementazione. Un'autorità terza e indipendente rispetto a quelle esistenti avrebbe garantito maggiore efficacia, neutralità e imparzialità, oltre ad assicurare un monitoraggio effettivo degli impatti etico-giuridici-sociali. Inoltre, sotto il profilo della *governance*, al fine di garantire un approccio olistico nel governo dell'AI, sarebbe stato opportuno prevedere un meccanismo di stretto coordinamento con altre autorità esistenti che esercitano indiscutibilmente funzioni in relazione all'intelligenza artificiale, al fine di realizzare strategie condivise in modo orizzontale, evitando verticalizzazioni, come il Garante per la protezione dei dati personali e l'Autorità Garante per le garanzie nelle comunicazioni; si prevedono invece solo generici e poco chiari concetti di "coordinamento e collaborazione".

Pertanto, la direzione italiana che porta a un disegno di legge in materia disvela alcune criticità alla luce della sovranità regolatoria europea, che a livello sostanziale cerca di affermarsi costruendo un modello di umanesimo tecnologico, basato su paradigmi e strumenti innovativi, che è necessario esaminare.

4. Umanesimo tecnologico: principi e strumenti giuridici

Nell'evoluzione che caratterizza il volto e la forma del diritto, che si esplica in *hard law* e *soft law*, e in cui si disvela la volontà di sovranità regolatoria, nella regolazione europea sono contenuti strumenti e direzioni, che costituiscono paradigmi nuovi a livello sostanziale, permettendo di ravvisare un contenuto innovativo del diritto. Il legislatore europeo si mostra consapevole infatti che, al fine di costruire un solido

[44] Tra le criticità sollevate rileva la prevista dualità tra AgID e ACN potrebbe essere problematica a livello applicativo anche nei rapporti con l'AI Office europeo. La Commissione europea ha inoltre ricordato che gli Stati membri non possono imporre requisiti tecnici o di conformità aggiuntivi per i sistemi che rientrano nel perimetro del regolamento; sotto tale profilo sarà interessante esaminare come saranno agite le previste deleghe al Governo per poter valutare eventuali disallineamenti.

modello di governo della tecnologia, ispirato ai principi degli ordinamenti democratici, è necessario superare le problematiche esistenti relative al rapporto tra diritto e intelligenza artificiale e, altresì, al rapporto tra essere umano e macchina, che trovano fondamento nelle criticità legate al funzionamento stesso dei dati e degli algoritmi quali il significativo squilibrio tra le parti in gioco e l'opacità dei processi di gestione.
In premessa è necessario precisare che nel quadro regolatorio europeo muta significativamente il generale approccio sostanziale alle geometrie di potere, dal momento che si sceglie di regolare e influire sui diversi attori in gioco, giacché si intende promuovere lo sviluppo economico e la competitività, ma si pongono obblighi di diligenza e trasparenza a carico delle aziende private, come nel caso del *Digital Services Act* e del *Digital Markets Act*, al fine di proteggere in modo più efficace gli utenti, istituendo un quadro di regole in materia di trasparenza e responsabilità.
Nel contesto di tale mutato approccio, che mostra l'intenzione dell'Unione europea di porre limiti all'autonomia delle aziende *Big Tech* alla luce del concreto ruolo di potere assunto nella società digitale, il modello europeo di governo della tecnologia fa leva su alcuni nuovi strumenti, soluzioni e paradigmi, capaci di innovare profondamente i paradigmi tradizionali e minimizzare i rischi, riequilibrando le asimmetrie a favore della collettività e proteggendo in modo efficace la persona.
In particolare, l'approccio innovativo riguarda proprio, da una parte, il rapporto tra diritto e tecnologia e, dall'altra, il rapporto tra essere umano e tecnologia.

4.1. Il rapporto tra diritto e tecnologia: *legal protection by design* e approccio basato sul rischio

Nella regolazione europea il rapporto tra diritto e tecnologia muta e matura nella costruzione di un umanesimo tecnologico[45] e di una *governance* umanocentrica, capace di conferire piena centralità alla persona, per tutelare la quale è necessario un bilanciamento mobile tra diritti in una logica che mantenga la tecnologia strumento nelle mani dell'uomo; tale approccio è abbracciato dagli atti di *soft law* e *hard law* relativi all'intelligenza artificiale[46].
Il carattere antropocentrico del modello europeo emerge in modo evidente nell'introduzione dell'obbligo di valutazione dell'impatto sui diritti fondamentali (*Fundamental Rights Impact Assessment* – FRIA), che non grava su tutti gli operatori, ma specificamente sui *deployer* (utilizzatori professionali) di sistemi di IA ad alto rischio che siano enti pubblici o soggetti privati che forniscono servizi pubblici essenziali (es. istruzione, banche, assicurazioni).
A differenza delle valutazioni puramente tecniche, la valutazione dell'impatto sui diritti fondamentali impone una riflessione di natura giuridica prima dell'impiego del sistema. Il *deployer* è tenuto a identificare i gruppi di persone vulnerabili potenzialmente esposti a rischi di discriminazione; valutare l'impatto su

[45] Cfr. A. Punzi, *Difettività e giustizia aumentata. L'esperienza giuridica e la sfida dell'umanesimo digitale*, in *Ars Interpretandi*, 1, 2021, 113 ss.
[46] Sulle problematiche poste dalla regolazione e dagli aspetti giuridici dell'AI cfr., inter alia, E. Maestri (a cura di), *Introduzione e note di sintesi al Regolamento* (UE9 2024/1689 sull'intelligenza artificiale), Napoli, 2024; G. Pascuzzi, *Il diritto dell'era digitale*, 6° ed., , Bologna, 2025, 303 ss.; G. Sartor, *L'intelligenza artificiale e il diritto*, Torino, 2022; F. Casa, S. Gaetano, G. Pascali (a cura di), *Intelligenza artificiale: diritto, etica e democrazia*, Bologna, 2025; C. Novelli, F. Casolari, A. Rotolo, M. Taddeo, L. Floridi, *AI Risk Assessment: A Scenario-Based, Proportional Methodology for the AI Act*, in *Digital Society*, 3, 2024, 1-29.

diritti specifici, quali la dignità umana, la libertà di espressione o il diritto alla protezione dei dati personali; definire misure di mitigazione concrete nel caso in cui la valutazione evidenzi rischi elevati[47].
Questo meccanismo mostra come la sovranità digitale europea aspiri ad essere una sovranità dei diritti: l'innovazione tecnologica deve essere preventivamente validata attraverso la lente dei valori democratici. Il *Fundamental Rights Impact Assessment* agisce quindi come un ponte tra l'astrattezza e la generalità delle norme e la concretezza degli algoritmi, istituzionalizzando il principio di *accountability* e garantendo che l'efficienza tecnologica non avvenga mai ledendo la tutela della persona.
La *governance* umanocentrica, guidata da un approccio etico-giuridico, può essere realizzata e implementata sfruttando la tecnologia stessa e la relazione tra regole giuridiche e informatiche; in specifico il diritto può avvalersi della tecnica per garantire il suo rispetto[48]. L'approccio preventivo e proattivo prende forma in due profili sinergici e connessi: il *legal protection by design* e il *risk-based approach*.
L'incorporazione preventiva di principi etici e giuridici, norme e rimedi nella tecnologia stessa, ossia una *legal protection by design*, basata sull'*accountability*, presente fin dal regolamento europeo 2016/679 in materia di protezione dei dati personali, emerge negli atti europei dedicati a dati e algoritmi quale paradigma capace di risolvere o quanto meno minimizzare le problematiche afferenti al rapporto tra diritto e tecnologia[49]. Il diritto può avvalersi della tecnologia per garantire il suo rispetto, generando un modello di "diritto nella tecnica"; si tratta di un approccio proattivo, che tutela la persona fin dalla progettazione, per impostazione predefinita e per mezzo della valutazione d'impatto, capace di far leva sulla conformazione dei sistemi tecnologici e sulla sicurezza informatica, da una parte, e sulla responsabilizzazione e sulla consapevolezza dei soggetti, dall'altra, al fine di confinare le repressioni prevalentemente ad una effettiva e persuasiva tutela sanzionatoria successiva, senza reprimere eccessivamente *ex ante* la libera circolazione dei dati e lo sviluppo economico.
Tale prospettiva è abbracciata e ulteriormente sviluppata nell'*Artificial Intelligence Act*, che si basa sul *risk-based approach*, approccio proattivo basato sul rischio coadiuvato anche da un correlato sistema sanzionatorio: il rischio viene sottoposto a una categorizzazione preventiva, distinguendo quattro categorie (rischio inaccettabile, alto, basso, minimo), cui si collega una regolazione differenziata.
Il sistema di responsabilità distingue in base alle categorie di operatori di sistemi di intelligenza artificiale ad alto rischio o meno, motivando la responsabilità dell'operatore con il fatto che sta controllando un rischio associato alla specifica tecnologia; tale modello proporzionato al rischio è consapevole dell'eterogeneità delle soluzioni tecnologiche di intelligenza artificiale, in cui può variare sensibilmente l'autonomia che le connota, la trasparenza che è possibile garantire e il controllo umano esercitabile.
L'approccio basato sul rischio, che mira a una protezione preventiva, capace di ridurre o eliminare la probabilità stessa che possano verificarsi violazioni, ha la capacità di raggiungere un equilibrio nella tutela dei diversi interessi in gioco, dal momento che è teso alla tutela dei diritti della persona, ma è anche orientato alla crescita economica, giacché il produttore di intelligenza artificiale può considerare la *legal compliance*

[47] Art. 27, reg. UE 2024/1689.
[48] In merito al rapporto tra etica e diritto in materia di AI cfr., inter alia, F.H. LLANO-ALONSO, *L'etica dell'intelligenza artificiale nel quadro giuridico dell'Unione europea*, in *Ragion Pratica*, 2, 2021, 327 ss.; M. CATANZARITI, *Etica "artificiale": un nuovo modello regolatorio?*, in *Ars Interpretandi*, 1, 2021, 165 ss.
[49] Si tratta dei principi data protection by design e by default, cui si affianca il data protection impact assessment (artt. 25 e 35, reg. UE 2016/679).

come un costo di produzione che entra nel costo economico della propria attività piuttosto che affrontare l'alea di dover eventualmente rispondere di possibili violazioni[50].

L'approccio proattivo, che prende vita nel *legal protection by design* e nel *risk-based approach*, non è scevro da criticità: la "rigidità" ontologica del codice informatico può determinare il rischio di un'eccessiva semplificazione di concetti e clausole complesse, si scontra con la flessibilità necessaria al bilanciamento "mobile" idoneo a una tutela efficace dei diritti e va a sommarsi all'opacità linguistica, dal momento che il linguaggio è informatico e non è quello naturale delle norme giuridiche.

Risulta complesso anche individuare quali principi etici e giuridici incorporare in tecnologie come l'intelligenza artificiale per lo più destinate a utilizzi transnazionali, dato il pluralismo etico e giuridico che caratterizza i diversi blocchi geopolitici e i differenti ordinamenti[51].

Ulteriore significativa problematica di tale prospettiva si annida nel fatto che in tal modo il rispetto dei principi giuridici e l'equilibrio tra diritti sono di fatto delegati a coloro che sono chiamati a sviluppare le soluzioni tecnologiche e alle categorie di operatori nel mercato con conseguenti possibili problematiche ed effetti degni di attenta considerazione.

4.2. Il rapporto tra essere umano e tecnologia: supervisione umana e trasparenza algoritmica

Nella regolazione dedicata all'intelligenza artificiale, accanto al rapporto tra diritto e tecnologia che si caratterizza per *legal protection by design* e *risk-based approach*, emerge il rapporto tra uomo e macchina, in cui si stagliano altri due paradigmi innovativi collegati sinergicamente ai primi: l'esigenza di mantenere il controllo dell'uomo sulla tecnologia, al fine di farle mantenere il ruolo di strumento che ontologicamente le spetta senza rischiare derive tecnocratiche, e, proprio anche a tal fine, la necessaria trasparenza nel funzionamento della stessa, che ne consenta comprensione, eventuale contestazione e sindacato da parte di un utente e/o di un giudice.

La necessità di mantenere servente la tecnologia rispetto all'uomo è sottesa all'attenzione che viene tributata al rapporto tra uomo e macchina, che si declina nella supervisione umana e nella non esclusività della decisione algoritmica.

La supervisione umana, tesa a prevenire e minimizzare rischi e pericoli, risulta in linea con l'approccio proattivo e preventivo che connota il modello di *governance* della tecnologia. Al fine di costruire un modello antropocentrico, infatti, è necessario mantenere equilibrio tra uomo e macchina, attraverso la garanzia dello "*human in the loop*"; al riguardo rileva il fatto che l'intelligenza artificiale non è più soltanto un mezzo per realizzare azioni, ma sempre più spesso è essa stessa a prendere autonomamente decisioni significative per la persona umana, laddove impiegata per tali scopi.

A tal fine l'*Artificial Intelligence Act* prevede misure concrete per garantire una supervisione umana sostanziale ed evitare il rischio che la decisione umana possa essere "attratta" dai risultati della macchina; il problema, infatti, si annida nelle influenze e nella capacità "attrattiva" a livello pratico della soluzione offerta dalla macchina e nella correlata difficoltà per l'uomo di discostarsi con la propria valutazione da

[50] In merito sia consentito il rinvio a F. FAINI, *Intelligenza artificiale e regolazione giuridica: il ruolo del diritto nel rapporto tra uomo e macchina*, in *federalismi.it*, 2, 2023, 1-29.

[51] Supra, § 2.

quanto emerge dal sistema tecnologico, finendo per conformarsi al risultato suggerito: in tal modo, di conseguenza, si svuota sostanzialmente l'autonomia della decisione umana[52].

Al riguardo però occorre interrogarsi se e come sia possibile conciliare supervisione umana con i sistemi di *machine* e *deep learning*, capaci di apprendere autonomamente, che operano secondo processi decisionali opachi e non prevedibili, rispetto ai quali pertanto devono essere verificate le effettive possibilità per l'uomo di mettere in discussione i risultati cui perviene la macchina, controllare la correttezza dei dati e correggere gli *output*.

A ben vedere costituisce condizione necessaria per la supervisione umana l'*explainability*, dal momento che solo se l'essere umano è in grado di comprendere il modo in cui l'intelligenza artificiale decide, può operare un controllo effettivo e, se necessario, correggerne gli *output*. E non a caso comprensibilità e sindacabilità sono paradigmi su cui si concentra la regolazione europea. Si tratta di garantire una declinazione rafforzata della trasparenza quale trasparenza algoritmica capace di assicurare conoscibilità, comprensibilità e sindacabilità, rendendo gli algoritmi oggetto della piena cognizione e del sindacato dal parte dell'uomo: tale principio si traduce nel garantire non solo informazioni e accesso ai dati, ma anche la conoscenza della logica che governa gli algoritmi, accompagnata dalla consapevolezza in merito alle conseguenze e all'impatto sulla persona.

Nei confronti dell'intelligenza artificiale, pertanto, devono essere attribuiti e riconosciuti il diritto alla comprensibilità, capace di rendere consapevole l'interessato, e il diritto alla contestabilità, idoneo a consentire all'interessato, anche per mezzo di un giudice, di valutare e di sindacare la decisione a cui perviene la tecnologia. Tali diritti sono declinazioni del più ampio e profondo diritto del singolo di mantenere autonomia, autodeterminazione e libertà nei confronti della macchina, che a sua volta comporta l'esigenza di comprensione dei meccanismi di funzionamento e una sorta di correlato dovere di "spiegare" in capo all'intelligenza artificiale o, meglio, in capo a chi ne è responsabile[53].

In linea con l'umanesimo tecnologico l'*explainability* assicura che l'intelligenza artificiale rimanga strumentale rispetto a quella umana e la tecnologia mantenga la sua funzione "servente" rispetto all'uomo e alle sue decisioni, soprattutto laddove incida su diritti e libertà[54].

L'*Artificial Intelligence Act* mostra particolare attenzione per la trasparenza, prevedendo che siano fornite informazioni chiare e adeguate all'utente sia in caso di sistemi ad alto rischio, sia in caso di sistemi a basso rischio. Ogni sistema ad alto rischio deve essere disegnato e sviluppato in modo da assicurare un appropriato livello di trasparenza (*sufficiently transparent*)[55], mostrando in tal modo consapevolezza circa la necessità di misure proporzionate, senza imporre obblighi irrealizzabili alla luce delle caratteristiche tecnologiche, come tali destinati ad una concreta inefficacia; emerge con evidenza l'esigenza di sostenibilità del diritto.

Alla trasparenza algoritmica della soluzione tecnologica si accompagna la necessaria trasparenza da parte di chi governa gli algoritmi, coniugandosi pertanto l'esigenza di comprensibilità delle macchine con la

[52] Art. 14, reg. UE 2024/1689, che prevede misure atte a consentire a chi è affidata la supervisione umana di comprendere e interpretare il sistema di AI, intervenire sul funzionamento del sistema o interromperlo, decidere in modo autonomo se utilizzarlo e «restare consapevole della possibile tendenza a fare automaticamente affidamento o a fare eccessivo affidamento sull'output prodotto da un sistema di AI ad alto rischio ("distorsione dell'automazione")».

[53] Cfr. M. PALMIRANI, *Big Data e conoscenza*, in *Rivista di Filosofia del diritto*, 1, 2020, 73 ss.

[54] U. PAGALLO, *Algoritmi e conoscibilità*, in *Rivista di Filosofia del diritto*, 1, 2020, 93 ss.

[55] Art. 13, reg. UE 2024/1689.

necessità di una correlata trasparenza da parte degli uomini che le gestiscono; in tal senso rileva il *Digital Services Act*, che impone alle piattaforme doveri di trasparenza e diligenza in merito agli algoritmi utilizzati, essendo tenute ad obblighi di motivazione nei confronti degli utenti e alla valutazione dei rischi sistemici collegati[56].

Al riguardo, la natura degli algoritmi pone il problema se esista sempre una logica comprensibile, dato il funzionamento degli stessi e la conseguente possibile non intelligibilità secondo criteri logico-razionali. L'intelligenza artificiale, come esaminato[57], che poggia su inferenze e sull'approccio statistico e probabilistico, determina talvolta difficoltà di comprensione circa le motivazioni delle risposte fornite. Pertanto garantire trasparenza algoritmica può essere particolarmente complesso a fronte di una congenita opacità degli algoritmi[58]. Tale aspetto può risultare problematico in caso di utilizzo dell'intelligenza artificiale in ambiti giuridici quali l'amministrazione pubblica e la giustizia, dal momento che sia i provvedimenti amministrativi (art. 3, legge 241/1990) sia quelli giurisdizionali (art. 111, comma 6, Costituzione) devono essere necessariamente accompagnati da una motivazione.

Inoltre la declinazione dell'*explainability* non è un concetto unitario, ma si articola su diversi livelli di granularità. In primo luogo, occorre distinguere tra spiegabilità *ex ante* e *ex post*. La prima riguarda la trasparenza statica del sistema, che si basa sulla documentazione tecnica e sulle logiche del modello, mentre la seconda attiene alla capacità di fornire una motivazione specifica per una singola decisione automatizzata, fondamentale per assicurare i diritti[59]. A ciò si aggiunge la dicotomia tra spiegabilità locale, che spiega il risultato per un singolo *input* specifico e globale, che fornisce una panoramica sul funzionamento complessivo dell'intero modello e sull'importanza delle diverse variabili.

Il quadro regolatorio promuove strumenti di trasparenza consolidati come le *model cards*, che documentano metriche tecniche, dataset di addestramento e limiti del modello di base e le *system cards*, che si concentrano sull'intero sistema in cui il modello è inserito, descrivendo l'interazione con l'utente e il contesto applicativo reale.

Peraltro l'obbligo di trasparenza incontra un limite fisiologico nella riservatezza delle informazioni, inclusi i segreti commerciali e i diritti di proprietà intellettuale. La regolazione europea mira a un'*accountability* significativa senza imporre la *disclosure* del codice sorgente o degli algoritmi proprietari, che comporterebbe un rischio di espropriazione tecnologica, ma impone di fornire informazioni sufficienti a comprendere la logica e i rischi del sistema: l'*AI Act* tenta di risolvere questo conflitto attraverso una trasparenza "funzionale", garantendo così il diritto dell'interessato a una spiegazione senza compromettere la sovranità economica dei fornitori.

Sotto il profilo del rapporto tra uomo e intelligenza artificiale, la regolazione europea si mostra consapevole delle caratteristiche peculiari dell'intelligenza artificiale generativa e gradua diversamente gli obblighi correlati. L'AI Act introduce, infatti, una distinzione fondamentale tra sistemi *General-Purpose AI* – GPAI (l'applicazione finale) e modelli *General-Purpose AI* – GPAI (l'architettura di base come i *Large Language Models*), imponendo obblighi che riflettono la posizione degli attori nella catena del valore. Mentre i

[56] Artt. 4, 34 e 35, reg. UE 2022/2065.
[57] Supra, § 1.
[58] Cfr. G. FIORIGLIO, *La Società algoritmica fra opacità e spiegabilità: profili informatico-giuridici*, in *Ars Interpretandi*, 1, 2021, 53 ss.
[59] Art. 86, reg. UE 2024/1689.

sistemi GPAI sono soggetti alle regole basate sull'uso specifico, i modelli GPAI devono rispettare requisiti di base (trasparenza, documentazione tecnica e sintesi sul rispetto del diritto d'autore). Per i modelli che presentano un rischio sistemico (identificato dal parametro della potenza di calcolo superiore a una certa soglia), gli obblighi si fanno più stringenti[60].

Tuttavia, rimangono aperti significativi punti di frizione nell'implementazione pratica. In primo luogo, esiste una tensione tra il dovere di fornire documentazione tecnica dettagliata lungo la *supply chain* e la protezione del vantaggio competitivo: la valutazione del modello (*adversarial testing*) è ancora priva di standard metodologici univoci, nonostante l'adozione del *Code of Practice* nel 2025. Inoltre, l'uso di una soglia rigida relativa alla potenza di calcolo per definire il rischio sistemico è critica per la sua natura statica, che potrebbe non cogliere l'evoluzione dell'efficienza algoritmica.

5. Conclusioni: riflessioni filosofico-giuridiche

Gli strumenti che emergono dal *framework* europeo, teso ad esercitare una sovranità regolatoria in materia, permettono di affrontare le problematiche che affliggono il rapporto tra diritto e tecnologia e la relazione tra essere umano e intelligenza artificiale, giacché intervengono sulle asimmetrie e sull'opacità contribuendo a "svelare" preventivamente problemi, consentendo di progettare la soluzione tecnologica in modo diverso; trasparenza e sindacabilità permettono di intervenire anche *ex post*, laddove il problema non sia emerso *ex ante*, proteggendo la persona.

Il modello giuridico di umanesimo tecnologico, che prende forma nella regolazione europea in materia, si serve della tecnologia, che può previamente implementare principi etico-giuridici con un principio di *legal protection by design*, deve essere comprensibile e connotata dalla supervisione umana sostanziale, oltre a basarsi sul *risk-based approach* e sull'*accountability* da parte di chi governa la tecnologia.

Nello sforzo regolatorio europeo si assiste a un rinnovamento del diritto che passa da una costruzione di matrice sovranazionale e *multistakeholder* ed è guidato da un approccio filosofico-giuridico, orientato verso la tutela dei diritti. Gli Stati sono chiamati a produrre norme, riconoscendo il riferimento fornito da una pluralità di ulteriori sistemi statuali, sovranazionali o extrastatuali. In tale contesto è corretto individuare, come prevede la regolazione europea, organismi sovranazionali dedicati al governo della tecnologia dotati di un certo grado di indipendenza rispetto a quei poteri che possono avere svariati interessi a orientarlo verso specifiche direzioni (poteri pubblici e privati).

Del resto, il cambiamento pervasivo determinato dallo sviluppo delle tecnologie e dalla centralità assunta dall'intelligenza artificiale investe gli equilibri tra diritti, il bilanciamento tra interessi, il rapporto tra poteri e la tenuta dei principi democratici fondamentali, influenzando le direttrici del futuro. Le scelte filosofiche, etiche e giuridiche fondamentali del modello europeo di governo della tecnologia coinvolgono il ruolo della persona, le geometrie di potere, il rapporto tra pubblico e privato.

Quale forma del diritto è necessaria perché le norme possano essere effettive? Quale modello e quali strumenti sono adeguati al mutato contesto? Quali sfide deve affrontare un modello di sovranità regolatoria digitale?

Nel quadro europeo emerge la sfida di trovare un punto di caduta tra certezza/prevedibilità e flessibilità/adattabilità garantendo equilibrio tra tutela della persona e promozione dello sviluppo economico, al

[60] Art. 51 ss., reg. UE 2024/1689.

fine di costruire un diritto efficace e sostenibile; del resto compito del diritto è operare un ragionevole bilanciamento tra interessi diversi. Sotto tale profilo la sfida si articola nella concreta capacità di implementazione dell'umanesimo tecnologico, senza che possa trasformarsi in un freno o un rallentamento dello sviluppo tecnologico e assicurando la sostenibilità del modello stesso per piccole e medie imprese e più ampiamente per il mercato.

L'approccio, i paradigmi e gli strumenti di *legal protection by design*, *risk-based approach*, supervisione umana e trasparenza algoritmica toccano e ruotano sul rapporto tra diritto e tecnologia, da un lato, e su quello tra essere umano e tecnologia dall'altro, con il comune scopo di riuscire a controllare e governare l'intelligenza artificiale, mantenendola strumento servente nelle mani dell'uomo, al fine di tutelare diritti e valori europei.

L'Unione europea esprime la propria sovranità regolatoria con un rafforzamento dello spazio giuridico pubblico, perduto nell'asimmetria tra utenti e piattaforme, che passa dal ruolo di organismi indipendenti, capaci ontologicamente di trovare quell'equilibrio tra certezza e flessibilità e creare, di conseguenza, un diritto sostenibile. Di conseguenza la riflessione deve concentrarsi anche sull'opportunità al momento dell'implementazione di un'applicazione normativa congiunta e di una forte collaborazione tra gli organismi sovranazionali previsti dall'insieme di norme dedicate alla dimensione digitale, abbracciando il necessario approccio olistico conforme alla complessità digitale e all'intreccio tra interessi diversi.

Oggi è quanto mai necessario individuare il modello di governo della tecnologia cui siamo diretti, che vede una nuova relazione tra diritto e tecnologia capace di garantire prevedibilità e certezza, proteggendo in modo efficace l'uomo rispetto all'intelligenza artificiale con strumenti dotati di flessibilità e adattabilità: un diritto sostenibile, capace di tutelare la libertà dell'uomo grazie a un saggio equilibrio tra diritti e interessi, facendo leva su un approccio olistico, al cui centro porre l'essere umano.