

Collana Tiresia - Future Studies and Law

Regole e Intelligenza Artificiale

Francesco Cirillo e Federico Girelli (*a cura di*)

C. Apostolo, F. Cirillo, N. Colacino, C. Conti
G. D'Alessandro, F. Girelli, S. Lazzara, P. Macrì
C. Parente, A. Pirozzoli, F. Salvatore, T. Spasari
A. Sterpa, M. Varriale

Quaderno 3

Editoriale Scientifica



TIRESIA
FUTURE STUDIES AND LAW
QUADERNO N. 3

Collana di studi sulle trasformazioni del diritto
e sui mutamenti delle libertà costituzionali

diretta da
Alessandro Sterpa

Tiresia rappresenta il mito della veggenza nel limbo identitario tra il divino e l'umano. Predire il futuro grazie a Zeus pur avendo perso la vista per colpa di Era, che lo aveva punito per aver vissuto sia il piacere sessuale da uomo che da donna, rivelandone le profonde diversità agli esseri umani. Il piacere è suddiviso in dieci parti, l'uomo ne riconosce solo una, la donna nove. La cecità di Tiresia è la condizione essenziale per guardare al futuro. Si tratta di un mito che coglie bene le difficoltà della scienza, resa in parte cieca quando ha difficoltà ad usare i tradizionali sensi – ossia gli istituti per il diritto – per leggere le trasformazioni presenti e prevedere gli assetti futuri della società e dei rapporti giuridici. La collana ospiterà interventi tendenzialmente brevi, frutto della collaborazione degli studiosi in occasione di seminari e convegni, al fine di permettere l'accrescimento delle occasioni di contaminazione reciproca. Aperta e plurale, la collana è ispirata al rispetto dei principi e dei valori del costituzionalismo liberale che ad oggi si conferma come l'unico metodo di convivenza umana che ha garantito sviluppo e libertà all'individuo e alla comunità politica.

Comitato scientifico

Fabio Addis, Vincenzo Antonelli, Simone Barbareschi, Carla Bassu,
Alberto Maria Benedetti, Maria Esmeralda Bucalo,
Andrea Cardone, Federica Fabrizzi, Michele Francaviglia,
Andrea Genovese, Giovanni Guzzetta, Emma Imparato,
Ignazio Lagrotta, Andrea Longo, Entela Cukani,
Massimo Nunziata, Giovanna Pistorio, Anna Romagnuolo,
Giulio Maria Salerno, Fiammetta Salmoni, Federico Savastano,
Alessandro Sterpa, Gianluca Sgueo, Alberto Spinosa, Simona Tiribelli,
Andrea Vernata, Maurizio Viroli, Nicola Viceconte, Daniela Vitiello.

I contributi del volume sono stati sottoposti a referaggio

Segreteria scientifica

Anna Bettoni, Claudia Capasso, Francesco Cirillo,
Candida Conti, Victor Maximilian Hartl

Regole e Intelligenza Artificiale

a cura di

Francesco Cirillo e Federico Girelli

Editoriale Scientifica
Napoli

Il volume è stato stampato con i contributi dell'Università degli Studi della Tuscia - Dipartimento di Economia, Ingegneria, Società e Impresa (DEIM)

Proprietà letteraria riservata

© Copyright 2026 Editoriale Scientifica s.r.l.
via San Biagio dei Librai, 39 - 80138 Napoli
www.editorialescientifica.com info@editorialescientifica.com
ISBN 979-12-235-0777-3

Ricordando Anna Pirozzoli

INDICE

Avvertenza 9

In ricordo di Anna Pirozzoli 13

CRISTIANA APOSTOLO, *La sfida dell'IA nel Codice dei contratti, tra esigenze di efficienza e attualità delle garanzie* 17

FRANCESCO CIRILLO, *Alcuni limiti teorici della regolazione tecnoetica nell'AI Act* 45

NICOLA COLACINO, *Tecniche di normazione dell'intelligenza artificiale tra hard e soft law: una lettura nichilista* 83

CANDIDA CONTI, *AI nel rapporto di cura* 121

GIOVANNI D'ALESSANDRO, *La Constitutional AI di Anthropic tra normatività, legittimazione e limiti* 137

FEDERICO GIRELLI, *La tutela delle persone con disabilità al tempo dell'intelligenza artificiale* 167

SALVATORE LAZZARA, *Paradigmi organizzativi e di apprendimento in Agenzia delle Entrate. L'impatto della digitalizzazione* 199

PAOLO MACRÌ, *Indirizzi strategici e principi normativi sull'utilizzo dell'Intelligenza Artificiale nella Pubblica Amministrazione* 209

CARMEN PARENTE, *EPR tessile e IA: profili di raccordo alla luce dell'AI Act* 227

ANNA PIROZZOLI, *L'intelligenza artificiale e il diritto: AI ruled ed AI ruler* 249

FRANCESCA SALVATORE, *Intelligenza Artificiale e Codice dei Contratti Pubblici* 265

TOMMASO SPASARI, *Sanità digitale e diritto alla riservatezza dei dati sanitari nell'era dell'intelligenza artificiale* 295

ALESSANDRO STERPA, *L'"algo-libertà": i prodromi del "diritto all'innovazione"* 323

MASSIMO VARRIALE, *L'uso degli algoritmi nell'analisi del rischio relativa al contrasto degli illeciti fiscali* 337

Avvertenza

Il presente volume nasce da un progetto scientifico dedicato alle trasformazioni che l'intelligenza artificiale imprime alle categorie del diritto pubblico, costituzionale e amministrativo. La professoressa Anna Pirozzoli ha preso parte a quest'iniziativa sin dalla sua genesi, condividendone l'impostazione e contribuendo al confronto che ne ha accompagnato la costruzione: per questo, d'intesa con il professor Alessandro Sterpa, Direttore della Collana, il volume è dedicato a Lei. La sua prematura scomparsa ha dato al lavoro un significato che non avremmo voluto, ma che abbiamo ritenuto doveroso riconoscere.

I saggi raccolti affrontano l'intelligenza artificiale da prospettive diverse, nella convinzione che il fenomeno non possa essere ricondotto a un quadro di categorie giuridiche stabilmente definito. Una parte significativa della riflessione si concentra sull'IA come oggetto di regolazione: l'*AI Act* europeo, il rapporto tra *hard law* e *soft law*, le discipline nazionali e settoriali, le forme di *governance* pubblica e privata. In questa prospettiva emergono le tensioni tra innovazione, tutela dei diritti fondamentali, sicurezza, mercato e legittimazione democratica; tensioni che nessuna formula tecnica riesce a dissolvere e, forse, una certa retorica tende a mascherare.

Altrettanta attenzione è dedicata all'impiego dell'intelligenza artificiale nei poteri pubblici. I contributi sui contratti pubblici, sulla pubblica amministrazione, sugli enti locali, sulla fiscalità e sull'organizzazione amministrativa mostrano come l'automazione non incida soltanto sugli strumenti dell'azione pubblica, ma anche sulle sue garanzie: trasparenza, motivazione, responsabilità, con-

trollo umano, imputabilità della decisione vengono ricalibrate o sfidate dal cambiamento tecnologico. L'efficienza dei dispositivi appare così come una promessa importante, ma mai del tutto autosufficiente a legittimare ogni uso possibile; il diritto, con la sua consueta lentezza, è chiamato a stabilire quali segmenti della decisione possano essere affidati alla macchina e quali debbano restare presidiati da responsabilità umana, muovendo dalla consapevolezza che tanto le macchine quanto i cervelli sono inclini a condizionamenti ed errori, forse diversi ma non perciò incommensurabili.

Il volume dedica poi attenzione ai diritti fondamentali e alle situazioni di vulnerabilità: disabilità, sanità digitale, riservatezza dei dati, consenso informato, giustizia predittiva, rischi discriminatori connessi ai sistemi automatizzati. In questi ambiti l'intelligenza artificiale si presenta insieme come enorme opportunità e talora come fattore di rischio: può ampliare l'accesso a servizi, cure e strumenti di inclusione, ma può anche accentuare asimmetrie, opacità e dipendenze decisionali. In questo *trade-off*, rinunciare ai vantaggi per paura delle possibili derive appare insensato, ma obliterare i pericoli cedendo a un entusiasmo ingenuo sarebbe imprudente.

Sul piano teorico e costituzionale alcuni contributi indagano la pretesa di incorporare valori giuridici ed etici nei sistemi tecnici, la retorica o la pretesa di un'IA affidabile, l'etica nel *design*, la trasformazione del costituzionalismo nell'ambiente digitale e il crescente ruolo di soggetti privati nella definizione di regole, standard e architetture decisionali. In tale prospettiva l'intelligenza artificiale è un dispositivo che costringe a ripensare il modo in cui il diritto produce categorie, distribuisce responsabilità e istituisce e regola spazi di decisione pubblica.

Il contributo di Anna Pirozzoli – dedicato al rapporto tra intelligenza artificiale e diritto nella duplice prospettiva dell'IA regolata e dell'IA capace di incidere sui processi giuridici – occupa una po-

sizione particolarmente significativa all'interno del progetto. Intercetta alcune delle domande che attraversano l'intero volume: quali limiti debbano essere posti all'automazione del giudizio, quale spazio resti al convincimento umano, come evitare che l'efficienza predittiva si traduca in compressione delle garanzie, quale ruolo possa ancora svolgere il diritto quando la tecnica pretende di anticipare, classificare o orientare la decisione.

L'opera non intende offrire una ricostruzione definitiva della materia; sarebbe, del resto, una pretesa poco prudente in un campo nel quale le norme inseguono, più o meno efficacemente, tecnologie che, nel frattempo, hanno già cambiato cappello e indirizzo. Lo stesso strumentario definitorio del regolatore europeo appare talvolta troppo specioso, talaltra troppo generico.

Il volume intende proporre una mappa critica di problemi: l'intelligenza artificiale osservata nei luoghi in cui incontra il diritto vivente, dalle amministrazioni alla sanità, dalla fiscalità ai contratti pubblici, dalla tutela delle persone vulnerabili alla giustizia, fino alle nuove forme di normatività privata, offrendo così un contributo ad un dibattito ancora aperto, nel quale entusiasmo tecnologico e ansia regolatoria devono essere entrambi, e con pari rigore, sottoposti a continua verifica.

Francesco Cirillo e Federico Girelli

In ricordo di Anna Pirozzoli

di *Federico Girelli**

Alle esequie della Professoressa Anna Pirozzoli onestamente non ho avuto cuore di intervenire anche nella veste di Suo amico: sarei stato sopraffatto dall'emozione. Mi sono così limitato a dare lettura del messaggio di cordoglio pubblicato dal Direttivo della nostra associazione scientifica, l'Associazione Italiana dei Costituzionalisti. Sono quindi particolarmente grato ad Alessandro Sterpa, Direttore della *Collana Tiresia – Future Studies and Law*, per avermi dato l'opportunità di ricordare, pur sinteticamente, il profilo scientifico di Anna ed essersi fatto lui stesso promotore dell'idea di dedicare appunto a Lei questo volume, che accoglie il Suo ultimo scritto; e naturalmente sono altresì grato a Francesco Cirillo, la cui efficace operosità ha reso possibile la realizzazione di questa iniziativa scientifica, così come a tutti coloro che hanno inteso parteciparvi.

Questa volta il sentimento di amicizia che mi lega da anni ad Anna Pirozzoli, oltre che sostenermi, difficilmente non trasparirà.

Professore ordinario di Istituzioni di diritto pubblico, Direttore del Dipartimento di Scienze Politiche, Giuridiche, Sociologiche e Umanistiche, Rettore Vicario dell'Università degli Studi Niccolò Cusano di Roma: la posizione accademica di Anna è a tutti nota. Quando Le chiesi, in relazione ai ruoli che ricopriva, come volesse essere appellata a voce ed anche nei documenti da elaborare as-

* Professore associato di Diritto Costituzionale e Coordinatore del Corso di Laurea in Giurisprudenza nell'Università degli Studi Niccolò Cusano di Roma.

sieme, mi disse: «al maschile, così come questi ruoli sono denominati negli atti ufficiali dell'Ateneo; quello che conta è appunto la funzione che svolgi (e come la svolgi), non il sesso a cui appartieni». La mia amica Anna guardava alla sostanza e proprio Lei, donna, divenuta professore ordinario giovanissima, non poteva esser certo tacciata di maschilismo.

L'attività di ricerca di Anna ha interessato gli ambiti più diversi del diritto pubblico e costituzionale e si è protratta sino al Suo "ultimo respiro", come testimonia il contributo pubblicato in questo volume. Mi limiterò (sol per ragioni di spazio) a ricordare le Sue tre principali monografie: *La dignità dell'uomo. Geometrie costituzionali* (2012); *Il potere di influenza degli organi di garanzia costituzionale* (2013); *Il potere di influenza nel sistema costituzionale. Le "forze politiche" inaspettate* (2019).

La prima, sviluppo e approfondimento di un altro lavoro monografico del 2007 (*Il valore costituzionale della dignità.* Un'introduzione), affronta nella prospettiva del diritto costituzionale un tema, la dignità umana, assai complesso, la cui declinazione in termini rigorosamente giuridici è opera davvero ardua.

Le successive due monografie dedicate al «potere di influenza» sono il frutto di un cimento scientifico di sicuro rilievo teorico. L'Autrice studia la "forza politica" di organi tradizionalmente esclusi dal circuito dell'indirizzo politico e che forze politiche certamente non sono. Una scelta metodologica originale, che porta nuova luce sul modo d'intendere le dinamiche ordinamentali, grazie appunto al canone del «potere di influenza». Ricordo quando Anna mi regalò il volume del 2013 e mi disse quanto l'avesse impegnata concettualmente questo lavoro, su cui in realtà aveva iniziato a ragionare ancor prima di dedicarsi al tema capitale della dignità.

Ferma la riconoscenza che tutti dobbiamo al Professor Alberto

Clerici che generosamente assicura la continuità operativa della direzione del Dipartimento, confesso che tuttora nella veste di Coordinatore del Corso di Laurea in Giurisprudenza a volte mi capita di pensare «su questa questione devo sentire Anna»; non solo, quando arrivo all'Università mi viene naturale (più di qualche volta) guardare se nel parcheggio ci sia la Sua automobile. L'automobile non c'è più… e all'occorrenza approfitto della squisita disponibilità di Alberto.

Anna ha sempre ricoperto il Suo ruolo di Direttore del Dipartimento con efficienza e con grande garbo, così chi ha avuto la fortuna di collaborare con Lei si è ritrovato subito destinatario di un benefico… "potere di influenza".

La sfida dell'IA nel Codice dei contratti, tra esigenze di efficienza e attualità delle garanzie

di *Cristiana Apostolo**

SOMMARIO: 1. Il contesto normativo e giurisprudenziale in tema di amministrazione algoritmica. – 2. Prove tecniche di governo dell'i.a.: i principi contenuti negli artt. 19 e 30 del nuovo Codice dei contratti pubblici. – 3. Dai principi alle regole: i non detti del Codice dei contratti pubblici e le prospettive interpretative alla luce della nuova disciplina della l. 23 settembre 2025, n. 132. – 4. Alcune conclusioni problematiche sull'impiego dell'i.a. nelle pubbliche amministrazioni anche alla luce della l. 23 settembre 2025, n. 132.

1. Il contesto normativo e giurisprudenziale in tema di amministrazione algoritmica

Il dibattito intorno alla portata applicativa da attribuirsi all'informatica pubblica nella transizione dalla dimensione analogica a quella digitale degli apparati amministrativi è andato sviluppandosi a partire dagli anni '70 ed intorno a due direttrici, ancora attuali, laddove – in estrema sintesi – la prima intravedeva nella tecnologia uno strumento di ausilio, la seconda giungeva a riconoscerle invece un ruolo propriamente creativo e determinante nell'ambito del procedimento decisionale[1].

* Dottoranda in Società in mutamento: politiche, diritti e sicurezza, Università degli Studi della Tuscia.

[1] Per una ricostruzione del ruolo dell'informatica nelle istituzioni M.G. LOSANO, *La lunga marcia dell'informatica nelle istituzioni italiane*, in R. CAVALLO

Se, per un verso, valorizzando la natura del procedimento quale scansione dell'attività amministrativa, questo era apparso già da tempo facilmente automatizzabile, dal punto di vista della sua funzione (pervenire a una decisione ponderando interessi in conflitto) la riflessione in merito alla possibilità di sostituire l'attività decisionale umana con quella informatizzata è pervenuta ad esiti ancora non univoci ma comunque polarizzati sulla natura (vincolata o discrezionale) della scelta da assumere, escludendone in particolare l'uso allorché fossero coinvolti concetti giuridici *indeterminati*, a causa dello spazio valutativo riservato al decisore e della non univocità del suo esito[2].

PERIN e D.U. GALETTA, *Il diritto dell'amministrazione pubblica digitale*, Giappichelli, Torino, 2020, XXIII e ss.

[2] Sull'ammissibilità dell'atto automatizzato di natura vincolata, G. DUNI, *L'utilizzabilità delle tecniche elettroniche nell'emanazione degli atti e nei procedimenti amministrativi. Spunto per una teoria dell'atto emanato nella forma elettronica,* in *Rivista amministrativa della Repubblica italiana*, 1978, p. 407 ss. e per una ricostruzione del relativo dibattito in dottrina L. VIOLA, *L'intelligenza artificiale nel procedimento e nel processo amministrativo: lo stato dell'arte*, in «Foro amministrativo», 2018, p. 1602 ss.; A. MASUCCI, *L'atto amministrativo informatico. Primi lineamenti di una ricostruzione*, Jovene, Napoli, 1993, 19; ID., *Procedimento amministrativo e nuove tecnologie. Il procedimento amministrativo elettronico ad istanza di parte*, Giappichelli, Torino, 2011, 89 ss.; A. STERPA (a cura di), *L'ordine giuridico dell'algoritmo*, Napoli, 2024; F. MOROLLO, *Documento elettronico fra e-government e artificial intelligence (AI), in federalismi.it*, Focus TMT, 2015, n. 2, pp. 17-18; M.S. GIANNINI, *L'interpretazione dell'atto amministrativo e la teoria giuridica generale dell'interpretazione*, Giuffrè, Milano, 1939; A. ANGIULI, *Studi sulla discrezionalità amministrativa nel quando*, Cacucci, Bari, 1988; L. BENVENUTI, *La discrezionalità amministrativa*, Cedam, Padova, 1986; U. RUFFOLO (a cura di), *XXVI lezioni di diritto dell'intelligenza artificiale*, Torino, 2021; D. PALAZZO, *Intelligenza artificiale e scelte amministrative. La discrezionalità alla prova degli algoritmi*, Dialoghi di diritto amministrativo, Roma, 2024. Sempre sul tema, in ambito giurisprudenziale, cfr. Tar Lazio, Sez. III-*bis*, 22 marzo 2017, n. 3769 operava ancora una differenziazione in merito alla possibilità di adoperare strumenti informatici a seconda della natura vincolata o discrezionale dell'atto da assumersi,

Tale approccio ha tuttavia rivelato le sue criticità tanto sul piano teorico, a fronte dell'indeterminatezza dei rispettivi concetti di attività vincolata e discrezionale e dell'incertezza del relativo confine, quanto su quello pratico, allorché negli ultimi anni si sono resi disponibili strumenti in grado di *replicare* anche attività decisionali complesse[3].

evidenziando come l'attività vincolata fosse "*compatibile con la logica propria dell'elaboratore elettronico in quanto il software traduce gli elementi di fatto e i dati giuridici in linguaggio matematico dando vita a un ragionamento logico formalizzato che porta a una conclusione che sulla base dei dati iniziali è immutabile*", laddove invece nell'ambito dell'attività discrezionale l'Amministrazione avrebbe "*la possibilità di scelta dei mezzi da utilizzare ai fini della realizzazione dei fini determinati dalla legge*", giungendo tuttavia ad affermare che "*l'ammissibilità dell'elaborazione elettronica dell'atto amministrativo non è legata alla natura discrezionale o vincolata dell'atto quanto invece alla possibilità, che tuttavia è scientifica e non invece giuridica, di ricostruzione dell'iter logico sulla base del quale l'atto stesso possa essere emanato per mezzo di procedure automatizzate quanto al relativo contenuto dispositivo*". Sull'applicabilità dell'i.a. all'attività vincolata anche Cons. St., Sez. VI, 8 aprile 2019, n. 2770 secondo cui "*L'utilità di tale modalità operativa di gestione dell'interesse pubblico* (i.e. quella automatizzata) *è particolarmente evidente con riferimento a procedure seriali o standardizzate, implicanti l'elaborazione di ingenti quantità di istanze e caratterizzate dall'acquisizione di dati certi ed oggettivamente comprovabili e dall'assenza di ogni apprezzamento discrezionale*".

Introducendo il concetto di "deontologia della discrezionalità", G. GUARINO, *Atti e poteri amministrativi*, Giuffrè, 1994 p. 200 e ss. individua le regole di comportamento elaborate dalla giurisprudenza per l'esercizio dell'attività discrezionale, riconducibili a loro volta a tre principi cardine: "*la esatta e completa informazione, la logicità e la coerenza*", tra queste: "*l'istruttoria deve essere compiuta attraverso uffici competenti ed agenti adatti: in modo particolare non può essere affidata a persone che non diano garanzia di scrupolo, obiettività, imparzialità*".

[3] Sul tema dell'affidabilità della procedura decisionale fondata sul ragionamento dell'i.a. J. KROLL, J. HUEY, S. BAROCAS, E.W. FELTEN, J.R. REIDENBERG, D.G. ROBINSON, H. YU, *Accountable Algorithms*, in «*University of Pennsylvania Law Review*», 165, 2016, pp. 695 ss. Sul tema della logica a-causale seguita nel ragionamento dell'algoritmo F. FAINI, *Il diritto nella tecnica: tecnologie emergenti e nuove forme di regolazione*, in «Federalismi.it», 2020, p. 112. Sull'uso di algoritmi predittivi F. COSTANTINO, *Rischi e opportunità del ricorso delle amministra-*

Con riguardo alla disciplina positiva, l'introduzione della tecnologia per lo svolgimento delle funzioni amministrative è stata oggetto di una prima codificazione ad opera dell'art. 3 d.lgs. 12 febbraio 1993, n. 39 (in seguito superata per effetto dell'entrata in vigore del d.lgs. 7 marzo 2005, n. 82 dopo essere parzialmente confluita nell'art. 3-*bis* L. 7 agosto 1990, n. 241 ss.mm.ii.) il quale, al comma 1, individuava nella *predisposizione* tramite sistemi informativi automatizzati la regola per l'adozione degli atti di tutte le pubbliche amministrazioni, specificando tuttavia come l'emanazione di atti amministrativi mediante sistemi informatici o telematici dovesse essere accompagnata "*dall'indicazione della fonte e del responsabile dell'immissione, riproduzione, trasmissione o emanazione*"[4].

Senonché, l'individuazione di quella informatica alla stregua di modalità standard per la *predisposizione* dell'atto amministrativo, stante anche il richiamo (ad opera del secondo comma dell'art. 3 cit.) al ruolo del responsabile, non pareva militare in favore di un'ammissione *tout court* dello strumento tecnologico in tutte le fasi della produzione provvedimentale, ponendosi piuttosto in continuità con quella linea di pensiero 'conservativa' che intravedeva nello strumento informatico un mezzo al servizio dell'attività deci-

zioni alle predizioni dei big data, in Diritto pubblico, 2019, p. 43 ss. Sulle caratteristiche dell'atto amministrativo vincolato, F. FOLLIERI, *Decisione amministrativa e atto vincolato*, in «Federalismi.it», 2017, p. 7; Id., *Logica del sindacato di legittimità sul provvedimento amministrativo. Ragionamento giuridico e modalità di sindacato*, Milano-Padova, Wolters-Kluwer-Cedam, 2017, 364 e ss.; L. VIOLA, *Atti vincolati, hard cases e riparto di giurisdizione*, in «Federalismi.it», 2017, 14.

[4] L'art. 3-*bis* L. 7 agosto 1990, n. 241 ss.mm.ii., aggiunto con legge n. 15 del 2005 prevede che «*per conseguire maggiore efficienza nella loro attività, le amministrazioni pubbliche incentivano l'uso della telematica, nei rapporti interni, tra le diverse amministrazioni e tra queste e i privati*». Sul carattere programmatico della disposizione, F. CARDARELLI, *Uso della telematica*, in M.A. Sandulli (a cura di), *Codice dell'azione amministrativa*, II ed., Milano, 2017, 421 ss.

sionale, ma seguitando a considerare quest'ultima una prerogativa irrinunciabilmente umana[5].

A fronte delle crescenti esigenze di efficientamento dei processi decisionali della P.A., da mero supporto operativo l'automazione è andata non solo ad inserirsi nell'architettura del procedimento ma anche nel nucleo dell'attività decisionale sostituendone in concreto parte dei processi, specie laddove richiedenti attività complesse, poco gestibili dal punto di vista quantitativo, tendenzialmente ripetitive[6]. La materia dell'utilizzo di algoritmi (via via più 'intelligenti') nell'attività amministrativa ed il dibattito sulla compatibilità delle procedure e del relativo *output* rispetto alla disciplina in materia di procedimento amministrativo sono così giunti all'esame del Giudice Amministrativo, tanto con riguardo al tema della conoscibilità

[5] Sulle cd. "scatole nere" F. PASQUALE, *Black box Society. The Secret Algorithms That Control Money and Information*, Cambridge, Massachusetts, London, England, 2015; R. CALO, D. KEATS CITRON, *The Automated Administrative State: A Crisis of Legitimacy*, in «Emory Law Journal», vol. 70, no. 4, 2021, 797 ss., spec. 805 ss. Sulla cd. "riserva di umanità" G. GALLONE, *Riserva di umanità e funzioni amministrative. Indagine sui limiti dell'automazione decisionale tra procedimento e processo*, Wolters-Kluwer-Cedam, Padova, 2023.

Sul rapporto tra provvedimento amministrativo in forma automatizzata e principio di legalità S. CIVITARESE MATTEUCCI, *"Umano troppo umano". Decisioni amministrative automatizzate e principio di legalità*, *Diritto Pubblico*, n. 5 2019, p. 5 e ss. e sullo stesso tema M. FRANCAVIGLIA, *L'intelligenza artificiale nell'attività amministrativa: principi e garanzie costituzionali nel passaggio dalla regula agendi alla regola algoritmica*, in «Federalismi.it, n. 17 del 2024. Il richiamo al principio di legalità è contenuto anche in Cons. St., Sez. VI, 28 aprile 2023, n. 4297.

[6] La quota maggioritaria della produzione giurisprudenziale in tema di utilizzo di algoritmi la si rinviene in materia di pubblico impiego e segnatamente con riguardo ai provvedimenti dell'allora MIUR in materia assegnazioni delle sedi al personale docente (tra cui Tar Lazio, Sez. III-*bis*, 22 marzo 2017, n. 3769, Tar Lazio, Sez. III-*bis*, 10 settembre 2018, n. 9230, Tar. Lazio, Sez. III-*bis*, 13 settembre 2019, n. 10964, Cons. St., Sez. VI, n. 8742/2019 cit.).

degli algoritmi adoperati dalle P.A., quanto in ordine alla legittimità della decisione stessa in presenza di procedure automatizzate[7].

Dopo iniziali oscillazioni[8], il G.A. si è orientato in senso favorevole rispetto all'automatizzazione[9] dei procedimenti amministrativi (riconosciuta quale strumento di efficacia e neutralità della decisione), sussumendo le relative fattispecie disfunzionali ai fenomeni ordinari di patologia dell'azione amministrativa ed appuntando le relative decisioni su aspetti quali la correttezza degli accertamenti istruttori, la coerenza della motivazione del provvedimento ovvero e, soprattutto, il rispetto del canone di trasparenza, a sua volta da declinarsi *in subiecta materia* attraverso 'nuovi' corollari specifici, gli ormai noti principi della *conoscibilità*, *comprensibilità*, *non esclusività della decisione* e *non discriminazione algoritmica*[10].

Ferma la facoltà di ricorrere allo strumento telematico per l'assolvimento delle funzioni amministrative – ritenuta espressione della generale potestà di auto-organizzazione delle P.A. e ricavata, sul piano della disciplina positiva, rispettivamente dagli artt. 3-*bis*

[7] B. RAGANELLI, *Decisioni pubbliche e algoritmi: modelli alternativi di dialogo tra forme di intelligenza diverse nell'assunzione di decisioni amministrative*, in Federalismi», n. 22/2020, 242 ss.

[8] Una critica particolarmente aspra al fenomeno dell'automatizzazione del procedimento giungeva in particolare da Tar Lazio, Sez. III-*bis*,10 settembre 2018, n. 9230 in cui si afferma come "*le procedure informatiche finanche ove pervengano al loro maggior grado di precisione e addirittura di perfezione non possano mai soppiantare sostituendola davvero appieno l'attività cognitiva acquisitiva e di giudizio che solo un'istruttoria affidata a un funzionario persona fisica è in grado di svolgere*".

[9] In particolare, Cons. St., Sez. VI, 8 aprile 2019, n. 2270 e Cons. St., Sez. VI, 13 dicembre 2019, n. 8472 e, di recente, Cons. St., Sez. VI, 6 giugno 2025, n. 4929.

[10] Sulla distinzione tra "spiegabilità" e "trasparenza" dell'algoritmo si veda la risoluzione del Parlamento europeo del 12 febbraio 2019 su una politica industriale europea globale in materia di robotica e intelligenza artificiale" (2018/2088). Tar Lazio, 9230/2018 cit.

della Legge sul procedimento amministrativo (L. 7 agosto 1990, n. 241 ss.mm.ii.) e 12 del Codice dell'Amministrazione Digitale (d.lgs. 5 marzo 2005, n. 82)[11] – va tuttavia osservato come la produzione giurisprudenziale in materia di attività amministrativa algoritmica sia maturata in un contesto fattuale nel quale il contributo fornito dall'innovazione tecnologica non era ancora assimilabile, in termini di autonomia, adattabilità e impatto[12], a quello proprio dell'intelligenza artificiale nel senso oggi inteso, vertendo quei casi su provvedimenti che risultavano assunti sulla base di uno schema esecutivo dotato, sì, di un elevato grado di automazione, ma pur sempre caratterizzato da una logica binaria, tale da renderli l'esito meccanico e predeterminabile di una scelta operata a monte e del tutto umana (sulla base dello schema *if-then*).

Analogamente dicasi per il quadro normativo sul quale sono andati formandosi detti orientamenti, che non teneva conto (né avrebbe potuto) del livello di innovazione cui si sarebbe pervenuti per effetto della c.d. rivoluzione digitale[13] e che solo di recente è stato integrato da una compiuta definizione degli strumenti di intelligenza artificiale per effetto della disciplina del Reg. UE 2024/1689 (cd. A.I. Act)[14].

[11] L'art. 12 d.lgs. 5 marzo 2005, n. 82 (cd. Codice dell'Amministrazione Digitale) impone alle amministrazioni di organizzare autonomamente la propria attività facendo ricorso alle tecnologie dell'informazione e della comunicazione per la realizzazione degli obiettivi di "*efficienza, efficacia, economicità, imparzialità, trasparenza, semplificazione e partecipazione, nel rispetto dei principi di uguaglianza e di non discriminazione*".

[12] R. CAVALLO PERIN e I. ALBERTI, *Atti e procedimenti amministrativi digitali*, in R. CAVALLO PERIN e D.U. GALETTA (a cura di) il Diritto dell'Amministrazione Digitale (cit.).

[13] Sul punto anche Cons. St., Sez. VI, 13 dicembre 2019, n. 8472.

[14] Sul punto sostiene G. GALLONE, *Riserva di umanità e funzioni amministrative. Indagine sui limiti dell'automazione decisionale tra procedimento e processo* cit. come "*Pur se il G.D.P.R. è senza dubbio rivolto anche alle Pubbliche Ammini-*

In questo contesto, il presente lavoro si propone di osservare l'assetto attuale della disciplina in materia rilevando le criticità tuttora riferibili al rapporto tra l'automazione e lo schema previsto dalla Legge sul procedimento amministrativo, con particolare riguardo al profilo della responsabilità e delle garanzie procedimentali. Da ultimo, si intende osservare la materia anche alla luce del più recente intervento legislativo in materia di i.a. (la l. 23 settembre 2025, n. 132) limitatamente agli aspetti regolatori che investono direttamente l'azione della p.a. e il procedimento amministrativo.

2. Prove tecniche di governo dell'i.a.: i principi contenuti negli artt. 19 e 30 del nuovo Codice dei contratti pubblici

Le disposizioni di rango primario sull'utilizzo dell'i.a. nel ciclo di vita dei contratti pubblici hanno rappresentato un approdo certamente innovativo (se non rivoluzionario) nell'ambito della produzione normativa in tema di modelli decisionali e forme gestionali delle PP.AA. basati su tecnologie informatiche ed elettroniche (cd. *e-government*)[15].

strazioni non si deve, tuttavia, obliterare la circostanza che la sua disciplina nasce nell'ottica (resa esplicita dagli artt. 1 e 2 del Regolamento e dal suo Motivo n. 151) di assicurare tutela alla riservatezza individuale e che lo stesso intercetta il tema della gestione automatizzata dei dati unicamente in tale prospettiva e non in quella, assai differente, delle garanzie di una buona amministrazione. […] Ne consegue che, per ipotesi, ove lo svolgimento di funzioni am ministrative automatizzate non dovesse implicare il trattamento di dati personali non troverebbe applicazione l'art. 22, par.1 del G.D.P.R. e, più in generale, il Regolamento medesimo".

[15] Con tale termine, in estrema sintesi, si vuole indicare il processo di informatizzazione della pubblica amministrazione che, adoperando le parole della Comunicazione del 26 settembre 2003 della Commissione Europea, può essere definito come "*l'uso delle tecnologie dell'informazione e della comunicazione nelle*

Oltre a tentare di innovare l'apparato di principi che governano la materia (introducendo, in particolare, quello del *risultato* e quello della *fiducia*)[16], il Codice dei contratti pubblici (d.lgs. 31 marzo 2023, n. 36) ha operato una rivisitazione – in chiave pragmatica – del ruolo dei soggetti della contrattualistica pubblica, proponendo, sul fronte dell'Amministrazione, una ridefinizione in chiave 'manageriale' del ruolo del Responsabile Unico di Progetto e, in favore degli operatori privati, introducendo nuovi istituti di semplificazione (a titolo di esempio attraverso l'introduzione del principio *once only*).

Per quanto di interesse in questa sede, la prospettiva del Codice in favore della logica del risultato investe anche il tema della digitalizzazione, al duplice scopo di accrescere l'efficienza dei processi decisionali e di salvaguardare le garanzie per le parti coinvolte, in particolare affrontando criticità preesistenti, riconducibili alla frammentazione degli obblighi in materia, attraverso strumenti quali l'acquisizione e creazione di soli documenti nativi digitali (in ossequio al principio '*digital by default*' indicato dalla Commissio-

pubbliche amministrazioni, coniugato a modifiche organizzative ed all'acquisizione di nuove competenze al fine di migliorare i servizi pubblici ed i processi democratici e di rafforzare il sostegno alle politiche pubbliche" (Comunicazione della Commissione Europea del 26 settembre 2003 «*Il ruolo dell'e-governement per il futuro dell'Europa*») cfr. Cons. St., Sez. VI, 8 aprile 2019, n. 2270.

[16] Il riferimento è agli artt. 1 e 2 d.lgs. cit. a loro volta riconducibili (al di là della portata 'comunicativa' delle denominazioni) rispettivamente ai canoni tradizionali di buon andamento, efficacia, efficienza (art. 97 Cost e art. 1, L. 7 agosto 1990, n. 241) e ai principi di trasparenza, imparzialità, legalità e tutela dell'affidamento. Sui principi enunciati dal Codice F. SAITTA, *I principi generali del nuovo Codice dei contratti pubblici,* www.giustiziainsieme.it, 2023 e sul loro ruolo nell'interpretazione giurisdizionale L.R. PERFETTI, *Sul nuovo Codice dei contratti pubblici. In principio*, in *Urb. e app.*, 2023, 5 ss.

ne UE) e l'uso di piattaforme in grado di interagire con le banche dati esistenti[17].

Il tema dell'uso dell'I.A. viene affrontato agli artt. 19 e 30 del Codice, laddove quest'ultimo detta i principi da adoperarsi nell'approvvigionamento di tali sistemi e nel loro utilizzo nella procedura, mentre il primo, altrettanto significativamente, eleva addirittura il ricorso a procedure decisionali automatizzate a 'principio' in materia di digitalizzazione del ciclo di vita dei contratti.

Nel trattare la fase di acquisto e sviluppo delle soluzioni tecnologiche e quella del loro impiego nell'attività decisionale, l'art. 30 (d.lgs. cit.) sancisce, in relazione alla prima, l'obbligo di assicurare la disponibilità di ogni elemento (compreso il codice sorgente) che sia utile a comprenderne il funzionamento e di prevedere nei relativi atti di gara clausole che assicurino le prestazioni necessarie alla correzione di "*errori ed effetti indesiderati dell'automazione*". A ben vedere la garanzia della *correzione* degli errori costituisce una traduzione in chiave applicativa di quel *contributo umano* già valorizzato dalla giurisprudenza amministrativa allorché individuava, quale condizione di validità dell'attività algoritmica, l'apprestamento di misure atte proprio a "*controllare, validare o smentire la decisione automatizzata*" [18]. Sulla medesima base il G.A. aveva codificato il diritto del privato (ricavabile dall'art. 22 Reg. UE 2016/689) ad essere destinatario di decisioni "*non prive di coinvolgimento umano*" e, quali ulteriori corollari, il principio di *riferibili-*

[17] Ancora sul tema D.U. GALETTA "*Digitalizzazione, Intelligenza artificiale e Pubbliche Amministrazioni: il nuovo Codice dei contratti pubblici e le sfide che ci attendono*" in Federalismi.it secondo cui i principi dell'art. 19 devono leggersi "*nell'ottica specifica di quel diritto ad una buona amministrazione che è ora codificato dall'art. 41 della Carta dei diritti UE e che implica e presuppone anche il trattamento imparziale ed equo, oltre che entro un termine ragionevole, delle questioni che riguardano i destinatari dell'azione amministrativa*".

[18] Cons. St., sez. VI, 13 dicembre 2019, n. 8472.

tà della decisione all'autorità e all'organo competente ad adottarla (enucleato dalla Carta della Robotica del 2017) ed il *diritto alla conoscenza delle ragioni a base della decisione* (a sua volta riconosciuto dall'art. 42 della Carta Europea dei Diritti Fondamentali)[19].

Analogamente, l'obbligo di rendere disponibile ogni elemento idoneo a comprendere il funzionamento della soluzione tecnologica, trova corrispondenza in quella *piena accessibilità* riconosciuta dal G.A. al privato e nella sua soggezione al sindacato di legittimità in quanto consistente in una regola giuridica (espressa attraverso un linguaggio informatico)[20]; entrambi aspetti consequenziali alla

[19] Il Consiglio dell'Unione europea ha in particolare esortato ad "affrontare l'opacità, la complessità, la faziosità, un certo grado di imprevedibilità e un comportamento parzialmente autonomo di taluni sistemi di IA, onde garantirne la compatibilità con i diritti fondamentali e agevolare l'applicazione delle norme giuridiche" Consiglio dell'Unione Europea, Conclusioni della Presidenza - La Carta dei diritti fondamentali nel contesto dell'intelligenza artificiale e della trasformazione digitale, 11481/20, 2020. Con riferimento ai principi etici in materia A. CELOTTO, *Come regolare gli algoritmi. Il difficile bilanciamento fra scienza, etica e diritto*, in Analisi giuridica dell'economia, n. 1/2019, 1 ss.

[20] In Tar Lazio, Sez. III-*bis*, 22 marzo 2017, n. 3769 il G.A. chiarisce la distinzione tra forma elettronica ed elaborazione elettronica dell'atto amministrativo, la prima coincidendo con l'uso dello strumento informatico per la redazione dell'atto, la seconda invece riguardando la provenienza informatica del contenuto dell'atto stesso, essendo in questo caso la macchina a provvedere al reperimento, al collegamento e alla interrelazione tra norme e dati assumendo un ruolo strumentale rispetto all'atto amministrativo conclusivo. 3742/2017)

Sul tema della qualificazione giuridica dell'algoritmo, R. CAVALLO PERIN, D.U. GALETTA (a cura di), *Il diritto dell'amministrazione pubblica digitale* cit.; A. PAJNO, F. DONATI, A. PERRUCCI (a cura di), *Intelligenza artificiale e diritto: una rivoluzione?*, Bologna, 2022; L. TORCHIA, *Lo Stato digitale*, Il Mulino, Bologna, 2023, 111; E. CARLONI, *I principi della legalità algoritmica. Le decisioni automatizzate di fronte al giudice amministrativo*, in Dir. amm., n. 2/2020, 273 ss.; A. SIMONCINI, *Amministrazione digitale algoritmica. Il quadro costituzionale*, in R. CAVALLO PERIN - D.U. GALETTA (a cura di), *Il diritto dell'Amministrazione pubblica digitale*, Torino, 2020, 1 ss Sul principio di legalità anche F. CARDARELLI, *Ammi-*

qualificazione del software alla stregua di *atto amministrativo in forma elettronica*[21].

Con riguardo invece all'impiego di tali strumenti nell'ambito della fase decisionale dell'affidamento vengono positivizzati i principi di *conoscibilità*, *comprensibilità*, *non esclusività della decisione* e *non discriminazione algoritmiche*.

In merito alla *conoscibilità* la giurisprudenza aveva dato atto del riconoscimento nell'ordinamento di un *interesse conoscitivo* della persona, codificato dagli artt. 13, 14 e 15 del GDPR ed in virtù del quale l'Amministrazione rimane vincolata a garantire la verifica a valle degli esiti del procedimento anche a garanzia dell'imputabilità della scelta al titolare del potere autoritativo individuato in base al principio di legalità[22].

nistrazione digitale, trasparenza e principio di legalità, in *Dir. dell'informazione e dell'informatica*, n. 1/2015, 227 ss.

[21] Più di recente, nel riferirsi all'uso di strumenti informatici Cons. St., Sez. VI, 28 aprile 2023, n. 4297 il G.A. ha adoperato la locuzione (a quanto consta inedita) "strumento istruttorio"; in quella sede si afferma "*Come già evidenziato dalla giurisprudenza di questo Consiglio (cfr. ad es. sentenza n. 8472 del 2019) il ricorso a strumenti informatici nelle procedure amministrative costituisce una modalità agevolata di istruttoria, senza che il singolo strumento – per quanto qualificabile in termini di intelligenza artificiale – possa, da un lato, derogare alle regole normative ed ai criteri posti a presupposto della singola procedura e, dall'altro lato, essere sottratto alla trasparenza nonché alla imputabilità all'amministrazione procedente. Quest'ultima è chiamata a verificare la correttezza del funzionamento dello strumento istruttorio utilizzato e la relativa coerenza agli obiettivi ed alle regole dettate, in coerenza al principio di legalità, per l'esercizio del potere in esame. E nel caso di specie l'amministrazione ha correttamente verificato come l'applicazione fatta dallo strumento automatizzato sia stata coerente ai criteri predeterminati normativamente. I criteri non sono dettati dallo strumento informatico ma dalla norma; e la relativa attuazione, rimessa per ragioni di economia e di speditezza ad un sistema automatizzato, è soggetta alla verifica da parte del decisore amministrativo, cui è imputata la relativa scelta*".

[22] Cons. St., sez. VI, 8472/2019 cit. e e, di recente, Cons. St., Sez. VI, 4929/2025 cit.

Sul punto, il Legislatore delegato, nel declinare il concetto di *conoscibilità*, fa riferimento alla *comprensibilità del software* ed al diritto di ricevere informazioni 'significative' sulla sua logica di funzionamento, facendo propria l'accezione più penetrante del principio di trasparenza (cd. *reasoned transparency*), di talché a dover essere del tutto accessibile sarebbe l'intero *iter* sottostante alla decisione informatizzata, comprensivo dell'indicazione, rispettivamente, degli autori del procedimento, del meccanismo di decisione, dei dati selezionati dal *software* e delle priorità assegnate da quest'ultimo[23].

Infine, l'art. 30 recepisce il canone della '*non discriminazione*

[23] Cons. St., Sez. VI, 8 aprile 2019, n. 2270. Si veda sul punto anche Cons. St., Sez. Cons. Atti normativi, 18 ottobre 2019, n. 2636 che nel valutare un regolamento attuativo del d.L. 13 agosto 2011, n. 138 art. 16 co. 25 (che a sua volta stabiliva che "*A decorrere dal primo rinnovo dell'organo di revisione successivo alla data di entrata in vigore del presente decreto, i revisori dei conti degli enti locali sono scelti mediante estrazione da un elenco nel quale possono essere inseriti, a richiesta, i soggetti iscritti, a livello regionale, nel Registro dei revisori legali di cui al decreto legislativo 27 gennaio 2010, n. 39, nonché gli iscritti all'Ordine dei dottori commercialisti e degli esperti contabili. Con decreto del Ministro dell'interno, da adottare entro sessanta giorni dalla data di entrata in vigore della legge di conversione del presente decreto, sono stabiliti criteri per l'inserimento degli interessati nell'elenco di cui al primo periodo, nel rispetto dei seguenti princìpi: a) rapporto proporzionale tra anzianità di iscrizione negli albi e registri di cui al presente comma e popolazione di ciascun comune; b) previsione della necessità, ai fini dell'iscrizione nell'elenco di cui al presente comma, di aver in precedenza avanzato richiesta di svolgere la funzione nell'organo di revisione degli enti locali;*

c) possesso di specifica qualificazione professionale in materia di contabilità pubblica e gestione economica e finanziaria degli enti pubblici territoriali") sottolineava "*la necessità di precisare nel testo del regolamento l'organo ministeriale competente a definire e ad adottare tale algoritmo e il relativo atto di adozione, nonché la necessità che tale provvedimento rechi un'adeguata motivazione sulla logica, i contenuti e il funzionamento del nuovo algoritmo, in modo da rendere previamente conoscibile, secondo un principio di trasparenza, il criterio operativo della estrazione a sorte, che presiede alla selezione dei soggetti aspiranti*".

algoritmica' a sua volta corollario del diritto alla *decisione non priva di contributo umano* e traducibile anch'esso nell'obbligo di apprestare adeguati strumenti di correzione dei dati in ingresso (e in uscita)[24]. Sul punto il Codice sancisce un duplice vincolo: con riguardo alla fase di decisione si prevede che il titolare del *software* debba adottare ogni misura atta ad impedire effetti discriminatori riguardo agli operatori economici (co. 3, lett. c); in linea generale poi – rispetto, cioè a tutte le fattispecie nelle quali ricorra l'uso di procedure automatizzate nel ciclo di vita dei contratti – si puntua-

[24] Il Considerando 71 del Reg. UE 2016/689 prevede il diritto dell'interessato a "*non essere sottoposto a una decisione, che possa includere una misura, che valuti aspetti personali che lo riguardano, che sia basata unicamente su un trattamento automatizzato e che produca effetti giuridici che lo riguardano o incida in modo analogo significativamente sulla sua persona, quali il rifiuto automatico di una domanda di credito online o pratiche di assunzione elettronica senza interventi umani. Tale trattamento comprende la «profilazione», che consiste in una forma di trattamento automatizzato dei dati personali che valuta aspetti personali concernenti una persona fisica, in particolare al fine di analizzare o prevedere aspetti riguardanti il rendimento professionale, la situazione economica, la salute, le preferenze o gli interessi personali, l'affidabilità o il comportamento, l'ubicazione o gli spostamenti dell'interessato, ove ciò produca effetti giuridici che la riguardano o incida in modo analogo significativamente sulla sua persona.* […] *In ogni caso, tale trattamento dovrebbe essere subordinato a garanzie adeguate, che dovrebbero comprendere la specifica informazione all'interessato e il diritto di ottenere l'intervento umano, di esprimere la propria opinione, di ottenere una spiegazione della decisione conseguita dopo tale valutazione e di contestare la decisione.* […] *Al fine di garantire un trattamento corretto e trasparente* […] *è opportuno che il titolare del trattamento utilizzi procedure matematiche o statistiche appropriate per la profilazione, metta in atto misure tecniche e organizzative adeguate al fine di garantire, in particolare, che siano rettificati i fattori che comportano inesattezze dei dati e sia minimizzato il rischio di errori e al fine di garantire la sicurezza dei dati personali secondo una modalità che tenga conto dei potenziali rischi esistenti per gli interessi e i diritti dell'interessato e che impedisca tra l'altro effetti discriminatori nei confronti di persone fisiche sulla base della razza o dell'origine etnica, delle opinioni politiche, della religione o delle convinzioni personali, dell'appartenenza sindacale, dello status genetico, dello stato di salute o dell'orientamento sessuale, ovvero che comportano misure aventi tali effetti* […]".

lizza come ricada sull'Amministrazione l'onere di adottare ogni misura volta a minimizzare il rischio di errori e inesattezze nei dati nonché di prevenire effetti discriminatori riguardanti le persone fisiche (co. 4).

Poste tali *regulae agendi*, il quadro normativo va integrato con le disposizioni dell'art. 19 (d.lgs. cit.) il quale da un lato, in chiave ricognitiva e programmatica, detta ulteriori principi operanti in materia di digitalizzazione delle procedure di affidamento (si tratta dei canoni di neutralità tecnologica, trasparenza, sicurezza informatica, protezione dei dati personali, che a propria volta si assumono attuativi dei diritti di cittadinanza digitale) e, sotto diverso – forse più pregnante – profilo, al co. 7, pone l'attenzione sulla possibilità, da valutarsi in relazione al tipo di procedura, di ricorrere a procedure automatizzate "*nella valutazione delle offerte*", indicazione, quest'ultima, dalla quale può ricavarsi anche il superamento delle perplessità talora avanzate in ordine alla possibilità di ricorrere a meccanismi di I.A. nell'ambito di attività non strettamente meccaniche e che presentino un coefficiente di discrezionalità[25].

[25] Sulla natura solo tecnica – e non teorica – degli ostacoli all'uso dell'i.a. nell'attività discrezionale si esprimeva Cons. St., sez. VI, 13 dicembre 2019, n. 8472: "*Né vi sono ragioni di principio, ovvero concrete, per limitare l'utilizzo all'attività amministrativa vincolata piuttosto che discrezionale, entrambe espressione di attività autoritativa svolta nel perseguimento del pubblico interesse. In disparte la stessa sostenibilità a monte dell'attualità di una tale distinzione, atteso che ogni attività autoritativa comporta una fase quantomeno di accertamento e di verifica della scelta ai fini attribuiti dalla legge, se il ricorso agli strumenti informatici può apparire di più semplice utilizzo in relazione alla c.d. attività vincolata, nulla vieta che i medesimi fini predetti, perseguiti con il ricorso all'algoritmo informatico, possano perseguirsi anche in relazione ad attività connotata da ambiti di discrezionalità. Piuttosto, se nel caso dell'attività vincolata ben più rilevante, sia in termini quantitativi che qualitativi, potrà essere il ricorso a strumenti di automazione della raccolta e valutazione dei dati, anche l'esercizio di attività discrezionale, in specie tecnica, può in astratto beneficiare delle efficienze e, più in generale, dei vantaggi offerti dagli strumenti stessi*".

3. Dai principi alle regole: i non detti del Codice dei contratti pubblici e le prospettive interpretative alla luce della nuova disciplina della l. 23 settembre 2025, n. 132

I principi fatti propri dal Codice (conoscibilità, comprensibilità, non esclusività e non discriminazione algoritmica) e l'esplicito riferimento al *contributo umano* evidenziano come il Legislatore delegato abbia accolta la linea di pensiero che dà per certo e necessario il costante controllo umano sull'operato delle tecnologie di i.a. Nondimeno ciò che nella configurazione della norma risulta difficile stabilire è l'*ubi consistam* di tale controllo, quali siano gli obblighi concreti in capo agli attori della procedura (primo tra tutti, ovviamente, il RUP) e quali ne siano le ricadute in tema di responsabilità.

Per quanto attiene al ruolo del fornitore del *software*, nell'assetto delineato dall'art. 30, deve concludersi che l'Amministrazione conserva – né avrebbe potuto essere altrimenti – le proprie responsabilità verso i terzi per gli eventuali atti illegittimi che risultino adottati tramite lo strumento tecnologico, e ciò anche in forza di quel dovere di vigilanza e prevenzione sancito dal co. 4, ferma la possibilità di rivalersi nei confronti del fornitore per le conseguenze risarcitorie ove gli eventuali errori siano riconducibili all'inadempimento degli obblighi di assistenza e manutenzione (art. 30, co. 2, lett. b) in capo a quest'ultimo.

Con riguardo alle responsabilità del RUP va dato atto che, in effetti, le funzioni esplicitate – con enfasi – dall'art. 15 co. 5 d.lgs. cit.[26] rientravano nelle sue prerogative anche nell'assetto previgente

[26] Recita l'art. 15 co. 5 d.lgs. 36/2023 cit. "*Il RUP assicura il completamento dell'intervento pubblico nei termini previsti e nel rispetto degli obiettivi connessi al suo incarico, svolgendo tutte le attività indicate nell'allegato I.2, o che siano comunque necessarie, ove non di competenza di altri organi*".

e sono rinvenibili, per quanto attiene al rispetto del termine, già nella Legge sul procedimento amministrativo (art. 2 e 2-*bis*) e, con riguardo al "*rispetto degli obbiettivi connessi all'incarico*" (per la cui individuazione a propria volta occorrerà tenere conto anche del contenuto dell'All. I.2, recante l'elenco delle attività affidate al RUP) e nella disciplina generale in materia di responsabilità erariale (art. 1 e ss., L. 14 gennaio 1994, n. 20); identiche considerazioni valgono per l'enunciazione – contenuta al secondo comma dell'art. 2 d.lgs. cit. – delle fattispecie che darebbero luogo a colpa grave ai fini della responsabilità amministrativa ("*violazione di norme di diritto e degli auto-vincoli amministrativi, […] palese violazione di regole di prudenza, perizia e diligenza e l'omissione delle cautele, verifiche ed informazioni preventive normalmente richieste nell'attività amministrativa, in quanto esigibili nei confronti dell'agente pubblico in base alle specifiche competenze e in relazione al caso concreto*") trattandosi di un'elencazione di portata solo descrittiva, inidonea ad innovare il contenuto degli obblighi di legge ricadenti sulle parti coinvolte.

Per quanto di interesse in questa sede, immutato l'assetto e il contenuto della responsabilità amministrativa, ai soggetti coinvolti – e in particolare al RUP – potranno, sì, imputarsi le condotte riconducibili all'inadempimento dei rispettivi obblighi, incluso il controllo sulle decisioni automatizzate, ma – e in maniera piuttosto ovvia – soltanto laddove detto soggetto sia stato posto in concreto nelle condizioni di intervenire ovvero di avvedersi dell'errore compiuto dal *software*, diversamente permanendo *in toto* la relativa responsabilità in capo all'Amministrazione; l'effettivo rispetto dei vincoli previsti dall'art. 30 transita difatti per l'appostamento, ad opera dell'Amministrazione, di misure concrete per consentire l'esercizio del controllo umano sulla decisione automatizzata, obiettivo che potrebbe realizzarsi anche avvalendosi della 'nuova'

configurazione del RUP quale 'responsabile degli obiettivi' e della facoltà di nominare ulteriori responsabili del procedimento che potrebbero coadiuvarlo anche per l'attività di controllo, validazione e correzione della decisione automatizzata[27].

Alla stessa conclusione si perviene osservando il tema del 'controllo umano' nell'odierna cornice normativa, come innovata dalla l. 23 settembre 2025, n. 132.

L'art. 14 (dedicato all'uso dell'intelligenza artificiale nella pubblica amministrazione) postula difatti, da un lato, una funzione dell'i.a. 'strumentale e di supporto all'attività provvedimentale', dall'altro ribadisce l'autonomia e il potere decisionale della persona chiarendo che quest'ultima "resta l'unica responsabile dei provvedimenti e dei procedimenti in cui sia stata utilizzata l'intelligenza artificiale".

Da ultimo, in chiave critica, deve rilevarsi che la valutazione in ordine alla scelta di ricorrere o meno agli strumenti tecnologici (compresa l'i.a.), in mancanza di indicazioni da parte del Codice, risulta a tutt'oggi rimessa alla stazione appaltante, senonché la genericità delle locuzioni adoperate (l'art. 30 prevede l'uso dell'i.a. "*ove possibile*" e l'art. 19 co. 7 d.lgs. cit. "*in relazione al tipo di procedura*") e il conseguente – ampio – margine di valutazione attribuito all'Amministrazione sull'uso o meno della tecnologia potrebbero tradire in parte l'intento del Codice di favorire la massima digitalizzazione aprendo a una differenziazione incontrollata del

[27] Per quanto attiene a quest'ultimo, significativamente, le sue competenze, le modalità per la sua individuazione e le relative funzioni vengono disciplinate all'interno della parte del Codice dedicata ai principi (Titolo II, Parte I, Libro I) e, segnatamente, all'art. 15 che, al quinto comma, affida al RUP il ruolo di assicurare "*il completamento dell'intervento pubblico nei termini previsti e nel rispetto degli obiettivi connessi al suo incarico*".

modus procedendi di ogni singola stazione appaltante[28], laddove invece sarebbe auspicabile una tendenziale uniformità, a beneficio della trasparenza e della concorrenza.

4. Alcune conclusioni problematiche sull'impiego dell'i.a. nelle pubbliche amministrazioni anche alla luce della l. 23 settembre 2025, n. 132

È stato precedentemente osservato come la più recente normativa nazionale in materia di i.a. (l. 23 settembre 2025, n. 132) abbia regolato – sebbene in una prospettiva prevalentemente a carattere programmatico – la materia dell'utilizzo di detti strumenti da parte della pubblica amministrazione.

Va dato atto anzitutto che mediante tale intervento normativo si è introdotta una base giuridica per l'utilizzo su larga scala di tecnologie di i.a. da parte delle pubbliche amministrazioni, superando così le perplessità connesse all'esistenza di una disciplina autonoma e speciale (ma isolata) per l'uso dell'i.a. nelle procedure di affidamento.

L'art. 14 l. cit., nel disciplinare l'uso dell'i.a. nella pubblica amministrazione, si appunta su tre elementi principali.

In primo luogo si enuncia l'interesse pubblico che sottende all'introduzione (o al riconoscimento) della facoltà per le amministrazioni di avvalersi di tali strumenti, il quale coincide con l'incremento di efficienza dell'attività amministrativa (in chiave di rapidità nella definizione dei procedimenti, aumento della qualità e

[28] Sulle problematiche connesse alla frammentazione degli obblighi in materia di digitalizzazione dell'attività amministrativa E. CARLONI, *Algoritmi su carta. Politiche di digitalizzazione e trasformazione digitale delle amministrazioni*, in Diritto Pubblico, Il Mulino, 2019.

quantità dei servizi erogati) e ciò ribadendo le garanzie per gli interessati in ordine alla *conoscibilità* del funzionamento dei software e la *tracciabilità* del loro utilizzo.

In secondo luogo, si affronta – sempre in un'ottica ancorata ai principi – il tema della relazione fra il potere decisionale e tali strumenti, riconoscendo a quest'ultimi una funzione strettamente 'strumentale e di supporto' rispetto all'attività provvedimentale e seguitando a postulare l'autonomia della persona, individuata come la sola responsabile dei provvedimenti e dei procedimenti in cui venga adoperata l'i.a.

In particolare, l'enunciato secondo cui la persona *resta* l'unica responsabile sembra esprimere l'intento, da parte del legislatore, di enfatizzare la stabilità e la continuità del regime di imputabilità degli atti e, sotto diverso profilo, il riferimento alla persona stessa quale responsabile *dei provvedimenti e dei procedimenti* in cui sia stata adoperata l'i.a. vale in effetti a rimarcare come nessun segmento dell'attività che conduce all'adozione del provvedimento finale (tanto in fase istruttoria quanto in sede decisionale) possa essere separato dalla responsabilità del *decisore umano*, e ciò anche ove nel suo operato questi sia stato coadiuvato dall'i.a.

È dunque da tale contesto, ossia in un assetto di responsabilità e imputabilità degli atti volutamente immutato da parte del legislatore, che deve ricavarsi la portata del cd. 'controllo umano', il quale, peraltro, non può ad oggi che inscriversi nella struttura del procedimento amministrativo, come configurata dalla L. 241/90.

Se, per un verso, deve darsi atto della portata innovativa dell'apparato di principi elaborati dal G.A. e dell'intervenuta conferma degli stessi attraverso la loro trasposizione all'interno del Codice dei contratti e, dipoi, nella l. 23 settembre 2025, n. 132, sotto diverso profilo, l'entità dello sforzo interpretativo compiuto per 'mettere a sistema' l'attività in forma algoritmica contribuisce a

descrivere la distanza tra detti moduli operativi e lo schema configurato dalla Legge sul procedimento, inducendo ad approfondire la riflessione sugli effetti della scelta in sé di automatizzare segmenti anche rilevanti di attività amministrativa e sulla loro compatibilità con quello schema[29].

Il profilo in questione può essere validamente affrontato, oltreché muovendo dai principi generali, anche a partire dal dato normativo, incentrando l'analisi sulla compatibilità del ricorso a strumenti di i.a. nell'attività amministrativa rispetto al disposto dell'art. 6 della Legge 241/90, il quale postula (al co. 1 lett. *a-e*) l'attribuzione al responsabile del procedimento di funzioni essenziali quali la *valutazione* delle condizioni di ammissibilità, dei requisiti di legittimazione e dei presupposti rilevanti per l'emanazione del provvedimento, l'*accertamento* dei fatti, l'adozione delle mi-

[29] In sede giurisdizionale è stato giudicato assorbente il profilo della violazione del principio di trasparenza e, talora, dell'obbligo di motivazione, tra tutte si veda Cons. St., Sez. VI, n. 8472/2019 cit.

La distinzione tra algoritmo ed intelligenza artificiale sarebbe stata ripresa dal Cons. St., Sez. III, 25 novembre 2021, n. 7891 allorché quest'ultimo, pur non affrontando il tema dell'operatività di detti strumenti nell'ambito dell'attività amministrativa, bensì dovendo darne una definizione per valutare la correttezza della scelta operata da una stazione appaltante a favore dell'applicativo fornito da un certo offerente anziché di un altro, individuava l'elemento discriminante nella capacità del software di creare un sistema in grado di "*elaborare costantemente nuovi criteri di inferenza tra dati*" e di "*assumere decisioni efficienti sulla base di tali elaborazioni attraverso un apprendimento automatico*" (cd. *machine learning*). Va peraltro rilevato come lo stesso principio di trasparenza sia suscettibile di più di un livello operativo di applicazione: sul punto è stata rilevata da C. COGLIANESE, D. LEHR, *Transparency and Algorithmic Governance*, in «Administrative Law Review», n. 1/2019, 1 ss. la distinzione tra due estrinsecazioni potenzialmente applicabili del medesimo principio, rispettivamente la forma 'attenuata' della *fishbowl transparency* (in virtù della quale dovrebbe consentirsi al privato di percepire dall'esterno l'attività svolta dall'Amministrazione) e quella – più penetrante – della *reasoned transparency* (che implicherebbe invece una comprensibilità ragionata della struttura e del funzionamento dell'algoritmo).

sure per l'adeguato e sollecito svolgimento dell'istruttoria, la cura delle comunicazioni, l'adozione, ove competente, del provvedimento finale e, in via generale, la *conduzione dell'istruttoria* (co. 2).

Non solo. La puntuale individuazione di un responsabile del procedimento amministrativo costituisce essa stessa una traduzione in chiave operativa del medesimo principio di trasparenza (e dei suoi corollari) cui dovrebbe conformarsi la stessa attività algoritmica, tanto che tale individuazione è qualificata dalla Legge sul procedimento amministrativo (all'art. 29 co. 2-*bis*) alla stregua di obbligo *attinente ai livelli essenziali delle prestazioni concernenti i diritti civili e sociali di cui all'art. 117 co. 2 lett. m) Cost*. Dal tenore complessivo dell'art. 29 co. 2-*bis* sembra peraltro potersi concludere che, ad onta del richiamo alla sola "*individuazione del responsabile*", a ricondursi ai LEP non siano le disposizioni strettamente riferibili alla designazione di tale soggetto, quanto piuttosto all'individuazione di un soggetto responsabile dotato di *quelle* specifiche prerogative indicate dall'art. 6cit., le uniche idonee a garantire le prestazioni di partecipazione e di trasparenza che, nella logica ispiratrice della L. 241/90, costituivano il principale elemento di rottura con il passato[30].

[30] Che quest'ultime non appaiono sacrificabili alla luce del processo di informatizzazione del procedimento amministrativo, risulta confermato anche dal CAD, allorché, all'art. 41 co. 2-*ter*, lett. c) richiede che il fascicolo informatico rechi l'indicazione del responsabile del procedimento. Sul tema da G. GALLONE in *Riserva di umanità e funzioni amministrativa. Indagine sui limiti dell'automazione decisionale tra procedimento e processo*" cit. sostiene che "*La menzione nell'ambito del Codice dell'Amministrazione Digitale della figura del "responsabile del procedimento" rappresenta, così, la più plastica dimostrazione sul piano del diritto positivo interno che l'informatizzazione dell'azione amministrativa, anche nelle forme del l'automazione decisionale, non può mai tradursi nella sua completa disumanizzazione*".

Sul tema, ad avviso di chi scrive, non appaiono persuasive le conclusioni cui sembra addivenire la giurisprudenza amministrativa allorché si è espressa sulla

Siffatta ricostruzione non appare smentita, ma vieppiù confermata dalla più recente disciplina nazionale in tema di i.a. laddove, difatti, si valorizza l'utilizzo di tali strumenti all'interno di una cd. *dimensione antropocentrica* (art. 1, l. 23 settembre 2025, n. 132) e dunque alla stregua di una visione che conservi il *primato dell'uomo* da intendersi, a propria volta, sia come 'persona' (come ribadito anche dall'art. 14) che in chiave valoriale.

Onde meglio demarcare la portata precettiva della ridetta *di-*

"*non integrale applicabilità al procedimento informatico della disciplina normativa generale sul procedimento amministrativo*" Cons. St., sez. VI, 13 dicembre 2019, n. 8472, difettando nell'ordinamento una fonte normativa che consenta di derogare agli obblighi previsti dalla L. 241/90. Altrettanto controversa l'asserzione secondo cui "*il tema dei pericoli connessi allo strumento (informatico n.d.r.) non è ovviato dalla rigida e meccanica applicazione di tutte le minute regole procedimentali della legge n. 241 del 1990 (quali ad es. la comunicazione di avvio del procedimento sulla quale si appunta buona parte dell'atto di appello o il responsabile del procedimento che, con tutta evidenza, non può es sere una macchina in assenza di disposizioni espresse), dovendosi invece ritenere che la fondamentale esigenza di tutela posta dall'utilizzazione dello strumento informatico c.d. algoritmico sia la trasparenza*".

Sul tema della complessità dell'agire amministrativo, osservava G. GUARINO in *Atti e poteri amministrativi*, 1994, p. 183 e ss come "*la descrizione del procedimento offre in modo plastico una rappresentazione della complessità della disciplina cui è assoggettato il potere amministrativo. La complessità è resa più ardua dal fatto che ciascun potere, che concorre alla serie procedimentale è rigorosamente vincolato all'osservanza di una propria disciplina specifica. Va da sé che la composizione del procedimento non è casuale, ma risponde alla finalità di coinvolgere, rappresentare e valutare tutti gli interessi che in qualche grado siano connessi con il provvedimento da emanare e con la materia che il provvedimento stesso è diretto a disciplinare. […] Se la descrizione del procedimento è utile come raffigurazione di insieme dell'agire amministrativo (che è oggi di norma un agire non per atti, ma per procedimenti) la individuazione dei principi che regolano la serie procedimentale in quanto tale, indipendentemente dalla sua composizione, e che valgono quindi per ogni procedimento, assume un valore essenziale: perché a tali principi corrispondono precisi vincoli ed effetti sul piano positivo e perché solo in base a tali principi si potrà apprezzare, fuori di qualsiasi visione astratta, la reale e concreta incidenza dell'istituto procedimentale*".

mensione antropocentrica può porsi mente all'art. 3 l. cit. allorché nell'enucleare i principi generali applicabili *in subiecta materia*, al primo comma, individua quale limite allo sviluppo, all'adozione e all'utilizzo di sistemi di i.a. proprio il rispetto dei diritti fondamentali e delle libertà previste dalla Costituzione e dal diritto dell'UE, ossia a quei valori indivisibili e universali quali – tra gli altri – la dignità umana, la libertà, l'uguaglianza, la solidarietà, nonché a principi fondamentali la democrazia e lo stato di diritto, a loro volta individuati dalla Carta dei Diritti Fondamentali dell'Unione quali presupposti per costituire uno spazio di libertà, sicurezza e giustizia.

Muovendo da tali coordinate di carattere generale nonché dal riconoscimento, operato dall'art. 29 co. 2 *bis* l. 7 agosto 1990, n. 241, dell'attinenza dell'istituto del responsabile del procedimento e delle garanzie di cui è portatore ai LEP, potrebbe allora inferirsi che, nel quadro normativo odierno, la facoltà dell'Amministrazione di avvalersi dei sistemi informatizzati (art. 3-*bis* L. 7 agosto 1990, n. 241 ss.mm.ii.) per lo svolgimento di attività – istruttoria o decisionale – incontri un limite interno in quel segmento di funzioni (di cui all'elenco dell'art. 6 cit.) che nella configurazione del procedimento amministrativo sono state demandate al responsabile, quest'ultimo potendo considerarsi tale (e dunque rispondere) soltanto con riguardo agli atti dei quali abbia la completa titolarità: e per averne curato in via diretta l'elaborazione in tutte le sue fasi e quando, benché materialmente generati da un *software*, li abbia fatti propri mediante un'attività di controllo a carattere sostanziale.

In difetto di tali precondizioni e quindi allorché lo strumento tecnologico, per come configurato, non consenta al responsabile di esercitare un controllo effettivo sull'attività automatizzata, per un verso questi non potrebbe in concreto esserne chiamato a rispondere dal punto di vista erariale (sia per difetto, nella sua condotta,

dell'elemento soggettivo della colpa grave, sia potendo configurarsi un concorso dell'Amministrazione nella causazione del danno rilevante *ex* art. 1337 c.c.) e, sotto diverso – e più severo – profilo si determinerebbe un arretramento delle garanzie di cui l'istituto stesso del responsabile è custode.

In questo senso – si noti incidentalmente – un riscontro in ordine al *quantum* del procedimento suscettibile di essere affidato all'automazione potrebbe forse ricavarsi muovendo proprio dal concetto stesso di responsabilità, il quale a sua volta, nella configurazione datane dall'art. 28 Cost., presuppone come fatto fondante proprio il *compimento* (e dunque la riferibilità concreta) di atti in violazione di legge da parte dei funzionari e dipendenti dello Stato[31].

In ordine all'*ubi consistam* del controllo (evocato dalla giurisprudenza, fatto proprio dal Codice dei contratti e promosso a principio generale anche dalla più recente l. 23 settembre 2025, n. 132, all'art. 3 co. 3) sembra dunque doversi concludere che la partecipazione del responsabile al procedimento non possa risolversi in un mero *contributo*, occorrendo invece un suo pieno e diretto coinvolgimento.

Pur dando atto del favore espresso dalla giurisprudenza verso

[31] In tema di responsabilità erariale, per una disamina esaustiva dei presupposti, si veda Corte Cost., 17 luglio 2024, n. 132. In dottrina, sul tema e senza pretesa di completezza, v. A. MONORCHIO, L. MOTTURA, *Compendio di contabilità di Stato*, Cacucci, Bari, 2021, 453 e ss.

Osservava ancora G. GUARINO in *Atti e poteri amministrativi*, 1994, p. 145 come "*I poteri amministrativi possono essere attribuiti solo a 'pubbliche amministrazioni'*" e, nell'elencarne le caratteristiche, che queste "*si avvalgono* […] *di dipendenti legati da rapporto di pubblico impiego* […]; *il personale assume necessariamente la qualifica di pubblico ufficiale o di incaricato di pubblico servizio ed è assoggettato alle relative disciplina penalistiche; il personale è soggetto alle ulteriori responsabilità disciplinare, civile, amministrativo-patrimoniale e, in determinati casi, contabile*".

il ricorso a *soluzioni algoritmiche* ed anche delle potenzialità di tali strumenti in tema di imparzialità ed efficienza, non può quindi non rilevarsi come nel modello vigente di procedimento amministrativo taluni segmenti di quest'ultimo debbano comunque risultare affidati all'attività umana: applicando tali coordinate ricostruttive al piano operativo, i vincoli dell'art. 6 (e per essi i principi precedentemente richiamati ricavabili dalla visione *antropocentrica* fatta da ultimo propria dal Legislatore) potrebbero astrattamente considerarsi soddisfatti in presenza di uno strumento algoritmico configurato in guisa da salvaguardare il *contributo umano*, ma a condizione che questo rimanga prevalente e decisivo rispetto a quello del sistema informatico e che siano parimenti rispettati i principi incorporati nell'art. 30 d.lgs. cit. (conoscibilità, comprensibilità, non esclusività e non discriminazione algoritmica).

Un simile assetto se, per un verso, può ritenersi agevolmente raggiungibile in presenza di *strumenti algoritmici* deputati a svolgere attività predeterminate (nei parametri e nei risultati), appare ben più complesso da ottenere con riguardo agli usi verosimili di *software* di *intelligenza artificiale* di tipo generativo, rispetto ai quali, anche applicando le citate coordinate, oltre a riproporsi i temi dell'effettività del controllo umano, dell'imputabilità e della responsabilità amministrativa per la decisione automatizzata, emergerebbe con evidenza quello dell'incompatibilità fra il loro uso e l'architettura del procedimento, imponendo una nuova definizione normativa del ruolo del responsabile, e, in ogni caso, dell'istruttoria nel suo complesso[32].

[32] Nell'ambito della ridefinizione in chiave manageriale della figura del RUP rientra altresì l'introduzione della facoltà per le stazioni appaltanti di individuare autonomi modelli organizzativi che prevedano l'affiancamento di due ulteriori responsabili del procedimento per le fasi della procedura (programmazione, progettazione, esecuzione e affidamento), con conseguente ripartizione di compiti e

Va dato atto, infine, che se, per un verso, il nuovo Codice dei contratti e la più recente disciplina generale in tema di i.a. sono valsi a legittimare in via astratta il ricorso a tali strumenti per lo svolgimento dell'attività amministrativa, il superamento delle già indicate criticità tanto in tema di responsabilità quanto – in via speculare – con riguardo al rispetto in concreto dei principi dettati per l'istruttoria procedimentale *in subiecta materia*, non può che postulare una definizione puntuale delle misure per l'effettivo rispetto degli obblighi stabiliti a carico dell'Amministrazione.

responsabilità, ma con la permanenza in capo al RUP delle funzioni di indirizzo, coordinamento e supervisione.

Alcuni limiti teorici della regolazione tecnoetica nell'AI Act

di *Francesco Cirillo*[*]

SOMMARIO: 1. Introduzione: il problema dell'etica regolatoria. – 2. L'*AI Act* e l'ipotesi del suprematismo assiologico occidentale. – 3. La questione dell'informazione semantica. – 4. *Explainability* e semantica artificiale. – 5. L'aporia epistemologica dell'*ethics by design*. – 6. La natura *technology-driven* dei valori. – 7. Conclusioni.

1. Introduzione: il problema dell'etica regolatoria

La regolazione dell'intelligenza artificiale si pone, fin dalla sua formulazione, in una condizione di tensione irrisolta: quella tra la pretesa di orientare *ex ante* lo sviluppo tecnologico mediante valori predeterminati e la natura intrinsecamente adattiva, contestuale e imprevedibile dei processi di innovazione. Questa tensione non è accidentale, né del tutto risolvibile per via tecnica o definitoria, perché riflette un limite strutturale del modello regolatorio che il legislatore europeo ha scelto di adottare. L'*AI Act* (Reg. UE 2024/1689; d'ora in avanti anche AIA) assume che i valori possano operare come principi anteriori e sovraordinati rispetto alle configurazioni tecniche, capaci di guidarne dall'esterno la progettazione e l'evoluzione, e che il decisore politico sia nella posizione di fissarli con sufficiente stabilità e determinatezza *ex ante*[1]. È su questo

[*] Docente di Metodologia giuridica e di Istituzioni di Diritto pubblico, Università degli Studi Niccolò Cusano.

[1] Per la critica all'idea che il diritto possa dirigere dall'esterno sistemi complessi tramite prescrizioni sostanziali, v. G. TEUBNER, *Substantive and Reflexive Elements in Modern Law*, in «Law & Society Review», 17(2), 1983, 239–286;

assunto che poggia l'impianto della regolazione europea, e sono proprio le sue fondamenta teoriche a meritare un esame critico[2].

La tesi è che tale impostazione presenti limiti teorici rilevanti, riconducibili a una serie di aporie tra loro connesse. Non è, in senso forte, l'etica a determinare la tecnologia; sono piuttosto le condizioni tecniche, materiali e operative a rendere possibili, a plasmare e a trasformare i sistemi di valori entro cui le pratiche sociali e giuridiche acquistano significato. In questo quadro, il diritto che pretenda di governare la tecnologia senza interrogarsi sui propri presupposti epistemici e sulle dinamiche effettive dell'innovazione rischia di ridursi a un dispositivo prevalentemente simbolico[3].

L'argomentazione complessiva si articola in quattro passaggi.

Il primo riguarda il fondamento teorico della regolazione europea, che fa ricorso a nozioni quali "intelligenza artificiale affidabile" o "antropocentrica" e presuppone la possibilità di incorporare valori astratti direttamente nei processi di progettazione tecnica[4]. È il modello, ormai ricorrente, della *compliance by design*, con-

quanto al limite strutturale della regolazione diretta dei sottosistemi sociali e tecnici, *Id.*, *Introduction to Autopoietic Law*, in *Id.*, *Autopoietic Law. A New Approach to Law and Society*, de Gruyter, Berlin–New York, 1988, 1 ss. Si accoglie qui la "verosimiglianza" della tesi dell'autopoiesi (o la sua plausibilità quantitativa), come "dispositivo strettamente euristico" in ragione del fatto che i modelli di causalità lineare (tra obiettivo politico, intervento normativo ed effetto sociale) si dimostrino empiricamente inadeguati.

[2] Contro una normazione che assuma principi come "dati" e li separa dalle condizioni operative, v. M. NEVES, *Symbolic Constitutionalization*, Oxford University Press, Oxford, 2022, spec. cap. 1.

[3] La produzione di consenso e la quiescenza attraverso atti normativi prevalentemente simbolici è stata oggetto degli studi di M. EDELMAN, *The Symbolic Uses of Politics*, University of Illinois Press, Urbana, 1964; sul rischio che principi astratti svolgano funzioni di legittimazione più che di razionalizzazione, cfr. J. HABERMAS, *Legitimation Crisis*, Beacon Press, Boston, 1973.

[4] COMMISSIONE EUROPEA, *Ethics Guidelines for Trustworthy AI*, 8 aprile 2019; «the foundations for a "Good AI Society"» in L. FLORIDI *et al.*,

cepito per fornire una possibile rassicurazione giuridica e sociale, che però risulta epistemologicamente molto fragile[5]. Il modello europeo assume che i valori possano essere tradotti in specifiche operative e possano essere programmati come vincoli funzionali, trascurando che essi non si presentano come entità modulari o formalizzabili senza residui[6]. Invece, si argomenterà che i valori emergano piuttosto dall'uso delle tecnologie, dai rapporti di potere che esse riorganizzano e dalle dinamiche economiche e sociali in cui sono inserite.

A ben vedere, l'etica regolatoria tende ad assumere i tratti di un vero e proprio "giusnaturalismo tecnologico": l'idea che esistano valori universali che l'intelligenza artificiale dovrebbe incarnare nella sua struttura, o che potrebbero essere impressi nella tecnologia attraverso un atto di progettazione, mentre i sistemi di AI operano in realtà secondo logiche contingenti, adattive e fortemente dipendenti dal contesto[7].

AI4People—An Ethical Framework for a Good AI Society, in «Minds and Machines», 28, 2018, pp. 689–70.

[5] La *self-regulation* potrebbe rafforzare la *compliance*, ma non sostituire l'*enforcement* tradizionale secondo J. L. SHORT, M. W. TOFFEL, *Making Self-Regulation More than Merely Symbolic*, in «Administrative Science Quarterly», 55(3), 2010, 361–396. Per un'analisi empirica degli effetti dello stile di *enforcement* sulla fiducia dei cittadini nelle autorità regolatorie, cfr. S. GRIMMELIKHUIJSEN *et al.*, *Does Enforcement Style Influence Citizen Trust in Regulatory Agencies?*, in «Journal of Public Administration Research and Theory», 35(1), 2025, 29–44.

[6] Si intende l'ipotesi, qui criticata, che un valore possa essere integralmente tradotto in specifiche operative tali da escludere ogni ulteriore spazio interpretativo.

[7] Sul carattere non neutrale e non "naturale" dell'incorporazione dei valori nella tecnologia v. L. LESSIG, *Code: Version 2.0*, Basic Books, New York, 2006, 83 ss. ("Regulating by Code"); e già L. WINNER, *Do Artifacts Have Politics?*, in «Daedalus», 109(1), 1980, 121–136. Sui limiti epistemici della *compliance by design* e del *privacy by design* v. M. HILDEBRANDT, *Smart Technologies and the*

Il secondo passaggio vorrà definire il quadro teorico entro cui collocare questa critica, spostando l'analisi dal determinismo etico a una prospettiva di co-costituzione tra tecnologia e valori. La tecnologia non agisce come una forza cieca e lineare che determina meccanicamente la società, ma certamente come un fattore che apre e chiude possibilità, ristrutturando progressivamente le pratiche, le istituzioni e, da ultimo, anche le categorie giuridiche. Lo stesso diritto può essere inteso come una tecnica sociale che entra in relazione con altre tecnologie, adattandosi e trasformandosi insieme ad esse, senza poter ambire a incidere in modo causalmente lineare e unilaterale su una determinata tecnologia.

La storia fornisce esempi ricorrenti di questa dinamica: la stampa a caratteri mobili ha ridefinito i rapporti tra sapere, autorità e fede, con effetti anche sulla cultura e la prassi giuridiche; la ferrovia ha contribuito all'unificazione dei mercati e alla nascita di nuove forme di regolazione economica; Internet ha imposto una revisione profonda delle nozioni di giurisdizione, *privacy* e libertà di espressione. In ciascuno di questi casi, non sono stati tanto (e solo) i valori astratti a guidare lo sviluppo tecnologico, ma soprattutto le configurazioni tecniche concrete a riplasmare i valori stessi e, così, anche le categorie giuridiche[8].

End(s) of Law, Edward Elgar, Cheltenham, 2015; nonché *Ead.*, *The Artifcial Intelligence of European Union Law*, in «German Law Journal», 21(1), 74–79.; L. EDWARDS, M. VEALE, *Slave to the Algorithm?*, in «Duke Law and Technology Review», 16(1), 2017, 18 ss. Sulla dipendenza contestuale dei valori v. H. NISSENBAUM, *Privacy in Context. Technology, Policy, and the Integrity of Social Life*, Stanford University Press, Stanford CA, 2009, 1 ss.; sul problema dell'universalismo etico, *inter alia*, A. SEN, *Human Rights and Asian Values*, in «The New Republic», 217(2–3), 1997, 33 ss.

[8] La tecnologia non opera come strumento neutrale né come fattore deterministico, ma come elemento che media e ristruttura le possibilità di azione, percezione e regolazione, contribuendo alla co-costituzione delle categorie sociali e giuridiche. In questa prospettiva, gli artefatti tecnici partecipano attivamente alla de-

Il terzo passaggio introduce un profilo spesso sottovalutato nel dibattito sulla regolazione, per certi versi assimilabile a una forma di pensiero magico che attribuisce al linguaggio normativo una capacità trasformativa diretta sulla realtà, in questo caso sulla realtà tecnologica. Siamo ben lontani dall'intuizione austiniana sugli atti linguistici, che coglie un aspetto difficile da contestare: qui, nel contesto del diritto digitale europeo, non si tratta di *fare* qualcosa dicendo, ma di credere che il *dire* basti a *produrre* l'oggetto tecnologico che si vorrebbe regolare o, mediante le regole, modellare. Il linguaggio normativo viene così investito di una funzione causale impropria, come se la qualificazione giuridica potesse sostituire i processi materiali e decisionali che presiedono alla progettazione e all'uso dei sistemi tecnici[9].

finizione di ciò che è possibile fare e regolare, trasformando dall'interno i quadri normativi di riferimento (P.-P. VERBEEK, *What Things Do* [2000], Pennsylvania State University Press, University Park PA, 2005, in particolare 203 ss.; B. LATOUR, *Reassembling the Social. An Introduction to Actor-Network-Theory*, Oxford University Press, Oxford, 2005, 1 ss.). Ne consegue che il diritto non si limiterebbe a governare la tecnologia dall'esterno, ma si dovrà modificare in relazione alle configurazioni tecniche disponibili, operando come una tecnica sociale in coevoluzione con altre (volendo, anche secondo l'intuizione di N. LUHMANN, *Recht als soziales System* [1986], in «Zeitschrift für Rechtssoziologie», 20, 1999, 1-13, per cui il diritto sarebbe un sistema autopoietico chiuso, e cognitivamente aperto alle informazioni esterne).

[9] La performatività del linguaggio non sarebbe una proprietà intrinseca delle parole, ma dipenderebbe da condizioni istituzionali e sociali che ne rendono possibile l'efficacia. Come argomentato dalla teoria degli atti linguistici, *dire* non equivale a *fare* se non in presenza di specifiche *felicity conditions* (J.L. AUSTIN, *How to Do Things with Words* [1955], Oxford University Press, Oxford, 1962, 1 ss.; J. SEARLE, *Speech Acts: An Essay in the Philosophy of Language*, Cambridge University Press, Cambridge, 1969, 16 ss.). In ambito giuridico, la forza performativa del linguaggio normativo deriva dal riconoscimento istituzionale e dal capitale simbolico che lo sostiene, non dalla mera enunciazione, vale a dire, da ultimo, dall'effettività del potere coercitivo dell'autorità da cui l'atto linguistico proviene (P. BOURDIEU, *La force du droit. Eléments pour une sociologie du champ ju-*

L'idea che qualificare normativamente un sistema come "affidabile", "etico" o "antropocentrico" sia di per sé sufficiente a orientarne il funzionamento presuppone una fiducia eccessiva nel potere performativo delle parole, come se il precetto del decisore politico potesse produrre automaticamente la realtà tecnologica.

In questo afflato performativo viene tuttavia trascurato un elemento strutturale del funzionamento del diritto: il circuito ermeneutico. È noto alla comunità dei giuristi che tra l'enunciazione normativa, la sua interpretazione, l'implementazione tecnica e l'uso concreto dei sistemi si apra, infatti, uno spazio di mediazione che non può essere colmato da definizioni astratte o da clausole valoriali, ma solo attraversato mediante pratiche interpretative, scelte organizzative e decisioni tecniche e politiche situate[10].

Nel diritto le parole non operano come cause immediate, ma producono effetti solo attraverso istituzioni, procedure e contesti determinati. Trasporre valori nel linguaggio della regolazione tecnologica senza interrogarsi sulle condizioni della loro interpretazione e attuazione significa attribuire al testo normativo una capa-

ridique, in «Actes de la Recherche en Sciences Sociales», 64, 1986, 3 ss.). Attribuire al testo normativo una capacità causale diretta sulla realtà tecnologica significa, dunque, confondere la funzione prescrittiva della norma con una pretesa efficacia materiale immediata, ignorando lo spazio di mediazione interpretativa, tecnica e istituzionale che separa l'enunciazione dalla sua attuazione (su cui anche M. HILDEBRANDT, *Smart Technologies and the End(s) of Law*, cit.).

[10] Il significato delle norme giuridiche non è dato una volta per tutte nel testo, ma si costituisce attraverso un circuito ermeneutico che comprende interpretazione, applicazione e uso; l'applicazione non è un momento successivo e meccanico, ma parte costitutiva dell'interpretazione stessa (per tutti, H.-G. GADAMER, *Verità e metodo* [1960], Bompiani, Milano, 2000). In ambito tecnologico, questo scarto si accentua, poiché la traduzione della norma in specifiche tecniche introduce ulteriori mediazioni interpretative che trasformano il significato dell'enunciazione normativa (J. ESSER, *Vorverständnis und Methodenwahl in der Rechtsfindung*, Athenäum, Frankfurt a.M., 1970; e ancora diffusamente in L. LESSIG, *Code: Version 2.0*, cit.).

cità causale che esso raramente possiede. Per queste ragioni, l'etica regolatoria rischia di funzionare come una semantica rassicurante, ma inidonea a raggiungere un margine di effettività[11]: un insieme di etichette che promettono controllo e orientamento, che però rimangono indeterminate e ambigue finché non vengono inserite nel circuito decisionale, tecnico e sociale che ne determina il significato effettivo. Da questo punto di vista, la regolazione non governa il processo tecnologico in corso; ne sospende simbolicamente la complessità, demandandone la gestione a sedi opache o tecniche. È significativo, peraltro, che proprio il circuito decisionale istituito dall'*AI Act* risulti in larga misura avulso dalla logica "tradizionale" del diritto pubblico, sia sotto il profilo delle garanzie procedurali, sia sotto quello della responsabilità istituzionale[12].

Il quarto passaggio chiarisce il profilo politico della critica che si muove alle premesse teoriche ed epistemiche della regolazione europea. Ogni ordine politico incorpora inevitabilmente una propria dimensione etica: non esiste decisione normativa priva di pre-

[11] Neppure poi tanto rassicurante: D. TAFANI, *L'«etica» come specchietto per le allodole. Sistemi di intelligenza artificiale e violazioni dei diritti*, in «Bollettino telematico di filosofia politica», 2023, 11 ss.

[12] O. POLLICINO, *Regolazione e innovazione tecnologica nell'"ordinamento della rete"*, in «Rivista AIC», 2, 2025, 164 ss., osserva alcune dei problemi congeniti della normativa. Cfr. A. SIMONCINI, *L'algoritmo incostituzionale: intelligenza artificiale e il futuro delle libertà*, in «BioLaw Journal», 1, 2019, 63 ss. Un approfondimento di questo aspetto in F. CIRILLO, *Un diritto per l'Intelligenza Artificiale: quis, quid, quando, ubi, cur?*, in «Democrazia e Diritti Sociali», 2, 2024, 199 ss.

I tratti descritti – distanza dai modelli idealtipici di *rule of law* e di *Rechtsstaat*, indeterminatezza normativa tale da rendere la regola conoscibile solo *ex post*, contraddittorietà, mutevolezza, imposizione di obblighi di dubbia fattibilità e ricorso strutturale a strumenti di *soft law* – possono essere letti come manifestazioni di una condizione di *post-diritto*, intesa come trasformazione della forma giuridica più che come semplice patologia della normazione. In questo senso, cfr. A. SCHIAVELLO, *Postdiritto: una breve guida per i perplessi*, in «Rivista di filosofia del diritto», 2, 2023, 261 ss.

supposti valoriali. Sarebbe dunque fuorviante criticare l'AIA per il solo fatto di esprimere una determinata concezione del bene o della dignità. Il problema teorico rilevante non è tanto se la politica debba avere un'etica, ma come essa la eserciti e dove la collochi. La pretesa problematica dell'etica regolatoria europea non consiste nell'imporre valori – cosa che ogni politica a modo suo fa –, bensì nel tentativo di trasformare scelte politiche controverse sul bilanciamento tra valori in conflitto in requisiti tecnici apparentemente necessari, sottraendole ai contrasti interpretativi, alla revisione giurisdizionale o al controllo democratico[13].

Inoltre, che il diritto possa e, in taluni casi, debba intervenire *ex post* per correggere effetti socialmente indesiderati delle tecnologie è difficilmente contestabile, e anzi spesso auspicabile. Ciò che appare invece problematico è la pretesa di imporre *ex ante* valori asseritamente universali, come se la tecnologia fosse poi in grado di incarnarli automaticamente[14]. Questo modello è, da un lato, inefficace, poiché la tecnologia evolve più rapidamente dei dispositivi

[13] La trasformazione di scelte politiche controverse in requisiti tecnici apparentemente necessari costituisce una forma di depoliticizzazione, che occulta il conflitto sotto l'apparenza della neutralità. Come mostrato dalla teoria politica, il politico riguarda decisioni fondamentali e conflittuali che non possono essere risolte tecnicamente (C. SCHMITT, *Il concetto di "politico"*, il Mulino, Bologna, 1972, 101 ss.). La riduzione di tali decisioni a problemi di razionalità tecnica opera come ideologia, mascherando scelte normative contingenti come necessità oggettive (già in J. HABERMAS, *Technik und Wissenschaft als »Ideologie«* [1068], Suhrkamp, Berlin, 2024). In questa prospettiva, la tecnicizzazione dell'etica e del bilanciamento tra valori produce una forma di *post-politics*, in cui il disaccordo viene sostituito dalla gestione esperta e amministrativa (la *post-political vision* su cui C. MOUFFE, *On the Political*, Routledge, London, 2005, 64 ss.). Poiché i conflitti tra valori sono strutturali e ineliminabili, nessun criterio tecnico può risolverli senza una decisione politica responsabile (I. BERLIN, *Due concetti di libertà* [1958], in *Id.*, *Libertà*, a cura di H. HARDY, Feltrinelli, Milano, 2025, 243 ss.).

[14] Cfr. A. PUNZI, *I Segni del tempo. Il ritorno della consuetudine nell'era della complessità*, in «Teoria e Critica della Regolazione Sociale», 2, 2021, 333 ss.

normativi che pretendono di anticiparla; dall'altro, ingannevole, perché tende a presentare come necessità tecnica ciò che è in realtà il risultato di scelte politiche contingenti, in relazione alle quali l'elemento della decisione rischia di restare opaco. La regolazione etica europea rivela il proprio carattere non universale, ma eminentemente politico, presentandosi proprio come un dispositivo di auto-legittimazione istituzionale che, nel tentativo di proporre un "modello europeo" di *governance* tecnologica, rischia così di mascherare decisioni controverse sotto l'apparenza di vincoli necessitati[15].

Se un determinato dispositivo tecnologico debba privilegiare la tutela della proprietà intellettuale o l'accesso all'informazione, la protezione delle minoranze o la libertà di espressione, la verità istituzionale o il libero mercato delle idee, il buon costume o la libertà sessuale non sono questioni suscettibili di soluzione mediante la progettazione informatica, né possono essere risolte affidando al codice il compito di "inverare" diritti fondamentali astrattamente definiti, o di istituire un ambiente digitale fondato sulla dignità umana, intesa in un senso del tutto indeterminato. Al contrario, le tecnologie riattivano il conflitto in forme storicamente e materialmente nuove: il rapporto tra *privacy* e informazione assume configurazioni diverse nella stampa tradizionale e nei *social network*; la tutela della proprietà intellettuale si trasforma radicalmente nell'orizzonte dei modelli linguistici di grandi dimensioni; la respon-

[15] Le tecnologie non risolvono i conflitti valoriali, ma li riconfigurano in forme storicamente e materialmente nuove. I valori emergono dai contesti d'uso e dalle *affordances* tecnologiche, richiedendo continue rinegoziazioni normative (*funditus* in H. NISSENBAUM, *Privacy in Context*, cit.). Ancora una volta, proprio perché i conflitti tra valori sono ineludibili, ogni tentativo di neutralizzarli attraverso soluzioni tecniche costituisce già una scelta politica (se si vuole, con la critica dell'*agonistic model of democracy* di C. MOUFFE, *The Democratic Paradox*, Verso, London, 2000, 80 ss.).

sabilità civile e amministrativa per i sinistri stradali muta profondamente nel contesto dei veicoli a guida autonoma. In ciascuno di questi casi, la tecnologia ridefinisce i termini di un conflitto che resta aperto in nuove declinazioni, rendendo illusoria ogni pretesa di neutralizzazione preventiva del *trade-off* tra valori mediante categorie etiche formalizzate[16].

Per queste ragioni occorre spostare il *focus* dal modo in cui incorporare i valori astratti nei sistemi tecnologici, vale a dire dal *design* etico, al riconoscimento del ruolo co-costitutivo che le tecnologie svolgono nella formazione e nella trasformazione dei valori stessi. Assumere questo rovesciamento di prospettiva consente di affrontare in modo più realistico le sfide poste dall'intelligenza artificiale e di interrogarsi su un diritto capace di operare efficacemente in un contesto in cui la tecnologia non è un semplice oggetto di regolazione, ma una co-condizione della normatività che si va definendo. Ciò sarebbe valso in generale per il passato, ma vale in particolare per le sfide del diritto nel mondo contemporaneo, dove ci si confronta in modo sempre più diretto con il fattore tecnologico, anche alla luce dell'accelerazione dei processi di innovazione caratteristici dell'età contemporanea.

[16] La distinzione tra intervento normativo *ex post* e regolazione *ex ante* è centrale nella teoria della regolazione delle tecnologie cd. emergenti. La regolazione anticipatoria presuppone capacità predittive che il diritto non possiede e tende a cristallizzare ipotesi contingenti sul futuro tecnologico (C. SUNSTEIN, *Il diritto della paura. Oltre il principio di precauzione* [2005], Il Mulino, Bologna, 2010; cfr. anche D. GUSTON, *Understanding "Anticipatory Governance"*, in «Social Studies of Science», 44(2), 2014, 218 ss., il quale scommette sulla capacità di intervenire su tecnologie emergenti prima che esse si cristallizzino). Il cosiddetto *Collingridge dilemma* mostra che, quando una tecnologia è ancora malleabile, non se ne conoscono ancora gli effetti, mentre quando gli effetti sono noti è troppo tardi per orientarla normativamente. Inoltre, la velocità dell'innovazione tecnologica rende strutturalmente obsoleti i tentativi di incorporare valori in requisiti *ex ante* («code changes faster than law», secondo L. LESSIG, *Code: Version 2.0*, cit.).

Si tratta, quindi, di una posizione consapevolmente critica, ma proprio per questo utile, soprattutto in un contesto in cui l'entusiasmo generale per le politiche regolatorie europee rischia di ridurre la riflessione teorica a un accompagnamento dell'agenda istituzionale, anziché a un autentico esercizio di scrutinio concettuale.

2. L'*AI Act* come manifestazione del suprematismo assiologico occidentale?

La regolazione europea dell'intelligenza artificiale non si limita a enunciare principi etici ritenuti genericamente condivisibili, ma può essere letta, in chiave critica, come una forma di "suprematismo assiologico": il tentativo di elevare valori storicamente situati e culturalmente connotati a parametri normativi presentati come universali. Un tentativo piuttosto incoerente, tanto più se affiancato dalla specificazione che i valori universali di cui si fa carico la regolazione sarebbero una caratteristica delle tradizioni giuridiche europee.

In questa prospettiva, l'AIA si presenta come l'esito di una capacità autonoma di orientare lo sviluppo tecnologico. Nondimeno, esso rischia di essere percepito più come una risposta politica alla fragilità strutturale dell'Unione europea nel contesto tecnologico globale. È proprio tale asimmetria che sembra indurre le istituzioni europee a fare dell'etica non solo un criterio di regolazione, ma anche uno strumento di legittimazione e di proiezione normativa[17].

[17] L'espressione 'suprematismo assiologico' è utilizzata in senso descrittivo e non morale per indicare il processo mediante il quale valori storicamente e culturalmente situati vengono presentati come parametri normativi universali e necessari, sottraendoli alla contestazione politica. In questo senso, l'universalismo non

Un primo elemento a sostegno di questa lettura riguarda la posizione dell'Europa nello scenario dell'innovazione tecnologica. Negli ultimi decenni, l'Unione si è progressivamente marginalizzata rispetto ai principali poli di sviluppo dell'intelligenza artificiale, in particolare agli Stati Uniti e alla Cina. In assenza di attori comparabili ai grandi *stakeholder* globali, una parte rilevante della capacità di intervento europea si è concentrata sulla produzione normativa, nella convinzione che la forza del mercato interno – il *Brussels Effect* – fosse sufficiente a proiettare gli standard europei oltre i confini dell'Unione. Nel caso dell'*AI Act*, tuttavia, come già accaduto per altre normative digitali, tale effetto appare tutt'altro che garantito. Diversi ordinamenti, tra cui Regno Unito, Giappone, Australia e Singapore, hanno adottato approcci più flessibili e maggiormente orientati alla promozione dell'innovazione, prendendo le distanze dall'impostazione europea, che appare ai critici sia burocratica che inefficace. Ancora più significativo è il fatto che gli ecosistemi tecnologici più dinamici, in particolare quello statunitense e quello cinese, mostrino una limitata propensione ad adattarsi a parametri di conformità modellati sul regolatore europeo. L'idea che gli altri continenti saranno spontaneamente attratti dall'*ethos* europeo appare alquanto dubbia.

Ne consegue che l'*AI Act* sembri esprimere meno una supre-

elimina la contingenza dei valori, ma la rende opaca. Sul carattere "provinciale" dell'universalismo europeo v. D. CHAKRABARTY, *Provincializing Europe. Postcolonial Thought and Historical Difference*, Princeton University Press, Princeton NJ, 2000, 16 ss.; sulla funzione di potere simbolico dell'universalizzazione normativa v. W. MIGNOLO, *The Darker Side of Western Modernity*, Duke University Press, Durham, 2011, 54 ss.; cfr. anche E. FERRI, *The Myth of Western Civilization*, Nova Science, Hauppauge NY, 2020. Nel caso europeo, tale dinamica si intreccia con una strategia di compensazione della debolezza tecnologica attraverso la proiezione normativa, come osservato da A. BRADFORD, *The Brussels Effect*, Oxford University Press, Oxford, 2020, 23 ss., su cui torneremo.

mazia normativa effettiva che un tentativo politico di compensazione, affidato alla regolazione, di una debolezza strutturale sul piano tecnologico[18].

Un secondo elemento concerne la struttura intrinsecamente assiologica della regolazione europea. L'AIA si fonda su categorie quali "intelligenza artificiale antropocentrica" e "intelligenza artificiale affidabile", presentate come valori di portata universale. Tali nozioni riflettono tuttavia una sintesi instabile di alcune tradizioni culturali, ora definite dalla centralità della persona o dalla tutela dei diritti fondamentali, così come declinate nelle esperienze politico-giuridiche europee. Non vi è nulla di intrinsecamente universale in queste categorie, soprattutto nella forma particolare che esse assumono all'interno della regolazione europea, spesso in contrasto con approcci differenti adottati in altri contesti.

Sotto questo profilo, la regolazione europea manifesta una dinamica che potrebbe anche essere letta essa stessa attraverso la categoria foucaultiana di *dispositivo*, che alcuni dei fautori della regolazione spenderebbero piuttosto o soltanto per la tecnologia regolata (e non per l'impalcatura che la regola): un insieme di pratiche, discorsi e istituzioni che si presentano come naturali o necessarie, ma che in realtà istituiscono e stabilizzano rapporti di potere. L'*AI*

[18] Le categorie di "antropocentrismo" e di "affidabilità" (*trustworthiness*) sono spesso presentate come ovvie e universalizzabili; in realtà rinviano a genealogie e presupposti culturali specifici. Sul carattere storicamente situato dell'umanesimo europeo e, quindi, dell'antropocentrismo come opzione valoriale (e non come dato acquisito o premessa incontestabile), v. R. BRAIDOTTI, *The Posthuman*, Polity Press, Cambridge, 2013, 13 ss.; Quanto a "affidabilità", essa non indica un contenuto neutro e autoevidente, ma dipende da assetti istituzionali e aspettative sociali variabili; v. O. O'NEILL, *A Question of Trust*, Cambridge University Press, Cambridge, 2002, 15 ss. Da ultimo, T. GRECO, *La legge della fiducia. Alle radici del diritto*, Laterza, Roma-Bari, 2021, offre una genealogia critica del paradigma antropologico "sfiduciario" che sottende il diritto moderno, mostrando come esso sia radicato in una specifica tradizione culturale e, perciò, non universale.

Act, proponendosi come guida etica globale, vorrebbe naturalizzare il proprio (incerto) orizzonte valoriale e a trasformarlo in uno standard tecnico, mascherando una scelta politica contingente sotto le sembianze di un'esigenza etica universale[19].

Un terzo profilo critico emerge dalla tensione interna tra l'enfasi assiologica dell'AIA e il principio, più volte richiamato, di neutralità tecnologica. Da un lato, l'Unione europea afferma che il legislatore non dovrebbe favorire soluzioni tecniche specifiche, limitandosi a fissare obiettivi generali (da perseguire *neutralmente*, vale a dire indipendentemente dai mezzi scelti). Dall'altro, nel momento in cui prescrive che i sistemi debbano essere "affidabili", "trasparenti" o "spiegabili", orienta di fatto lo sviluppo tecnologico verso determinati modelli e architetture. L'attenzione riservata alle "ontologie" e alla cosiddetta *semantic AI* ne è un esempio significativo: privilegiando strutture formali e linguistiche per garantire tracciabilità ed *explainability*, la regolazione finisce per restringere il campo delle soluzioni tecnicamente ammissibili, penalizzando approcci alternativi, come i modelli neurali puramente statistici, meno compatibili con una griglia semantica predeterminata[20].

Da questa prospettiva, il regolamento europeo può essere letto

[19] Anche volendo adottare la prospettiva dei critici della tecnologia e assumere in senso rigoroso la nozione foucaultiana di *dispositivo*, l'attenzione analitica deve spostarsi dalla tecnologia alla regolazione stessa. In Michel FOUCAULT, il *dispositif* è un insieme eterogeneo di discorsi, istituzioni e pratiche che produce regimi di verità e stabilizza rapporti di potere. Letto in questa chiave, l'*AI Act* non si limiterebbe a disciplinare l'intelligenza artificiale, ma contribuisce a costituire le categorie attraverso cui essa diviene conoscibile e governabile, definendo nozioni come "rischio" o "affidabilità" che non precedono la regolazione, ma ne sono l'effetto.

[20] La proclamata neutralità tecnologica della regolazione è difficilmente sostenibile. Come ha mostrato L. WINNER, gli artefatti incorporano sempre scelte (*lato sensu*) politiche, che diventano opache una volta tradotte in forma tecnica (*Do Artifacts Have Politics?*, cit.).

come un caso di razionalizzazione istituzionale *ex post*: il tentativo delle istituzioni di imporre ordine a processi tecnologici che seguono una dinamica in larga misura autonoma, proiettando su di essi una griglia di valori che risponde più a esigenze di legittimazione politica che alla reale logica dell'innovazione. Non è, in questo schema, la tecnologia a diventare più "affidabile" perché resa conforme a determinati standard; è piuttosto la regolazione a stabilire che cosa debba essere considerato affidabile, trasformando un'opzione politica controversa in un vincolo tecnico apparentemente necessario.

Una possibile obiezione sostiene che, pur riflettendo un orizzonte culturale particolare, l'*AI Act* assolverebbe comunque a una funzione imprescindibile di tutela dei diritti fondamentali in un contesto globale spesso privo di garanzie adeguate. L'argomento è tutt'altro che irrilevante, ma poggia su un presupposto teoricamente discutibile: che la protezione dei diritti richieda necessariamente l'imposizione *ex ante* di categorie etiche presentate come universali. Una regolazione maggiormente orientata agli effetti concreti delle tecnologie, piuttosto che alla loro pretesa conformità valoriale preventiva, apparirebbe più coerente con un orizzonte realmente condivisibile e più solida rispetto alle competenze giuridiche dell'Unione[21].

A ciò si aggiunga una distanza strutturale tra l'ambizione etica enunciata dalle istituzioni e la loro capacità effettiva di incidere sui contesti in cui i diritti risultano sistematicamente compromessi. In un quadro globale segnato da violazioni diffuse e persistenti, anche dell'ordine internazionale, gli strumenti giuridici disponibili – inclusi

[21] Sulla centralità degli strumenti *ex ante* (*risk management*, FRIA) e sui limiti della tutela effettiva dei diritti nell'*AI Act*, v. M. EBERS, *The AI Act Roller Coaster: The Evolution of Fundamental Rights Protection in the EU's Artificial Intelligence Regulation*, in «European Journal of Risk Regulation», 2025, 1 ss.

quelli europei – mostrano una capacità limitata di produrre effetti incisivi, stabili ed esportabili. È in questo scenario che l'enfasi sull'etica preventiva applicata alle tecnologie rischia di assumere, come detto, una funzione prevalentemente simbolica, quasi una forma di auto-rappresentazione normativa in un ambito emergente, là dove il costo politico dell'intervento risulta più contenuto. Sebbene questa dinamica sia già visibile nel GDPR, nel *Digital Services Act* e nel *Digital Markets Act*, anche l'*AI Act* può essere letto come una forma di politica commerciale indiretta, funzionale a selezionare gli attori ammessi al mercato europeo e a compensare, mediante barriere normative, una persistente debolezza industriale dell'Unione.

D'altra parte, il marcato investimento simbolico sull'etica dell'intelligenza artificiale, come pure sulla retorica delle generazioni future, a fronte di un atteggiamento ben più prudente – quando non ambiguo – rispetto a violazioni manifeste dei diritti umani e della legalità internazionale, finisce per indebolire la credibilità della pretesa universalistica della regolazione. L'etica invocata dall'*AI Act* appare così meno come un criterio coerente di giudizio normativo e più come un dispositivo retorico selettivo, applicato con rigore verso un oggetto tecnologico in larga parte ancora indeterminato, e con cautela verso ambiti in cui il costo politico dell'intervento risulterebbe immediato e concreto.

3. La questione dell'informazione semantica

Uno dei presupposti epistemici più problematici dell'*AI Act* riguarda l'idea che i dati possano fungere da "informazione semantica fattuale": elementi grezzi, neutrali e intrinsecamente dotati di significato, idonei a costituire un fondamento oggettivo per la regolazione etica dell'intelligenza artificiale. Tale impostazione, riconducibile

almeno in parte alla filosofia dell'informazione di Luciano Floridi, presuppone l'esistenza di unità informative portatrici di un contenuto semantico autonomo; unità che consentano in modo sufficientemente di stabilire una distinzione tra vero e falso e, su questa base, la costruzione di un'etica dell'infosfera[22].

Si sosterrà, invece, che questo presupposto sia epistemologicamente debole. La difficoltà non è dovuta all'opportunità di singole scelte tecniche o applicative, ma investe il modo stesso in cui il rapporto tra dato, significato e verità è concettualizzato sul piano teorico e poi trasposto sul piano normativo.

La teoria matematica dell'informazione, nella formulazione classica di Shannon, non stabilisce alcun nesso tra informazione e verità. L'informazione è definita come una variazione registrata all'interno di un sistema, rilevante unicamente in quanto riduce l'incertezza, a prescindere dal contenuto semantico o dall'aderenza alla realtà. Attribuire al dato un significato oggettivo e intrinsecamente veritativo equivarrebbe, dunque, a un'estensione indebita della teoria, caricandola di una funzione semantica ed etico-normativa che essa, nel modello di Shannon, non può avere.

Sul piano epistemologico, inoltre, la filosofia della scienza ha da tempo messo in discussione l'idea che esistano "dati puri". Ogni os-

[22] L'idea che i dati costituiscano informazione semanticamente significativa e veridica, idonea a fondare un'etica dell'infosfera, è centrale nella filosofia dell'informazione di Luciano Floridi, che definisce l'informazione semantica come dati ben formati, significativi e veri (L. FLORIDI, *Outline of a Theory of Strongly Semantic Information*, in «Minds and Machines», 2004, 197 ss.; *Id.*, *From Data to Semantic Information*, in *Philosophy of Information*, Oxford University Press, Oxford, 2011, 85 ss.). Tale presupposto è stato criticato sul piano epistemologico, poiché attribuisce ai dati una semanticità autonoma che prescinde da pratiche interpretative e criteri esterni (cfr., p.e., P. ADRIAANS, *A Critical Analysis of Floridi's Theory of Semantic Information*, in «Knowledge, Technology & Policy», 2010, 131 ss.)

servazione è mediata da schemi concettuali, strumenti tecnici e pratiche di classificazione. Da Kuhn a Quine, fino a Putnam, è emerso con chiarezza via via maggiore che i dati dipendano (non solo, ma anche) dai quadri teorici e dai contesti interpretativi: ciò che viene registrato come un mero "fatto" è sempre il risultato di una selezione[23]. In questo senso, parlare di *raw data* sarebbe un vero e proprio ossimoro: la produzione dei dati presuppone già scelte di rilevanza, di inquadramento e di significato[24], per cui se ci riferiamo ai *data*, essi non potranno essere *raw*.

Questa fragilità epistemica ha conseguenze decisive anche sul piano giuridico. Se i dati non sono neutrali né semanticamente autosufficienti, fondare su di essi un'etica *by design* produce un corto circuito concettuale: ciò che viene presentato come "fondamento oggettivo" della regolazione è in realtà il risultato di scelte interpretative precedenti. In questo modo, decisioni politiche e valoriali vengono occultate dietro una pretesa necessità epistemica, e categorie come "affidabilità", "equità" o "antropocentrismo" rischiano di essere trattate come proprietà tecniche dei sistemi, anziché come oggetto di deliberazione pubblica e giuridica. Richiedere – o

[23] L'idea di dati come entità puramente fattuali e pre-interpretative è incompatibile con la filosofia della scienza post-positivista. A partire da KUHN, fino a QUINE e PUTNAM, è emerso che l'osservazione è sempre mediata da schemi concettuali, strumenti e pratiche teoriche: ciò che viene registrato come "dato" è già il risultato di una selezione guidata da presupposti interpretativi. Non esistono dunque "fatti grezzi" indipendenti da quadri teorici di riferimento. T.S. KUHN, *The Structure of Scientific Revolutions* [1962], II ed., University of Chicago Press, Chicago, 1970, 125 ss.; W.V.O. QUINE, *Two Dogmas of Empiricism* [1951], in *From a Logical Point of View,* Harvard University Press, Cambridge MA, 1961, 20 ss.; H. PUTNAM, *Reason, Truth and History*, Cambridge University Press, New York, 1981.

[24] Per tutti, cfr. il volume di L. GITELMAN (a cura di), *"Raw Data" Is an Oxymoron*, MIT Press, Cambridge MA, 2013, 1 ss.; o G.C. BOWKER, *Memory Practices in the Sciences*, MIT Press, Cambridge MA, 2005, 183 ss.

anche solo auspicare – che gli algoritmi "dicano il vero" e non inducano in errore appare comprensibile e, sul piano pratico, forse desiderabile. Tuttavia, una simile pretesa risulta concettualmente inconferente, poiché proietta sui sistemi computazionali un criterio di verità che non appartiene al funzionamento statistico o formale degli algoritmi.

Si può obiettare che, pur non essendo neutrali, i dati mantengano una forma di resistenza oggettiva che consente di distinguere tra verità e falsità: essi sarebbero, se non fondamenti di verità, quantomeno vincoli alle opzioni interpretative. Ma questa osservazione non scioglie il nodo teorico. Che i dati impongano vincoli empirici non significa che contengano già la semantica necessaria per orientare una regolazione etica. Essi delimitano al più un campo di possibilità, ma non fissano i criteri di rilevanza, né definiscono il significato normativo delle informazioni prodotte.

Il punto epistemicamente più critico, però, riguarda la confusione, implicita in molte formulazioni regolatorie, tra semantica umana e semantica computazionale. Nella tradizione filosofica e linguistica, la semantica è inseparabile da significato, intenzionalità e verità (con o senza la 'v' maiuscola). Un enunciato è significativo perché rinvia a uno stato di cose nel mondo ed è vero se e solo se tale stato di cose sussiste. La verità, pur sempre mediata, resta un criterio extra-linguistico, ancorato al giudizio umano e al riferimento alla "realtà".

La cosiddetta *semantic AI* opera invece su un piano radicalmente diverso. La sua "semantica" consiste in mappature statistiche o formali che associano espressioni linguistiche a strutture interne al sistema, sulla base di correlazioni apprese o di ontologie predefinite. Non vi è intenzionalità comunicativa, né accesso allo stato di cose cui un enunciato rinvia.

Il sistema non "sa" se una proposizione sia vera; si limita a va-

lutare se essa sia compatibile con un *pattern* statistico o con una classificazione formale. Anche in applicazioni giuridicamente rilevanti, come l'analisi automatizzata di contratti o decisioni, ciò che viene prodotto non è un giudizio di verità, ma l'inclusione di un testo entro una categoria operativa[25].

Questa distinzione assume un rilievo ancora maggiore nel caso dei *Large Language Models*. Tali sistemi generano sequenze linguistiche plausibili sulla base di modelli probabilistici di predizione autoregressiva. L'*output* è ottimizzato per coerenza sintattica e contestuale, non per aderenza alla realtà – qualsiasi cosa essa sia. Le cosiddette "allucinazioni" non costituiscono perciò inspiegabili anomalie occasionali, ma effetti strutturali del funzionamento del modello, che deve sempre produrre un *output*, anche in assenza di una base fattuale adeguata. Potremmo dire che le allucinazioni siano anomalie per la nostra semantica, ma non per quella dei modelli. L'idea di "informazione semantica fattuale" si rivela, perciò, internamente ambigua: nessun sistema fondato su correlazioni testuali può garantire la dimensione veritativa (che si vorrebbe) propria del linguaggio umano[26].

Nemmeno il ricorso alle cosiddette ontologie formali consente di colmare questo divario. Le ontologie permettono di organizzare concetti e relazioni secondo schemi coerenti, ma non eliminano la

[25] Sulla distinzione tra manipolazione sintattica e comprensione semantica, nonché sull'assenza di intenzionalità e riferimento nei sistemi computazionali, v. J. SEARLE, *Minds, Brains, and Programs*, in «Behavioral and Brain Sciences», 3, 1980, 417 ss.; e, oltre ai contributi già citati, L. FLORIDI, *Is Semantic Information Meaningful Data?*, in «Philosophy and Phenomenological Research», 70, 2005, 351 ss.

[26] Sulla natura autoregressiva dei *Large Language Models*, sull'ottimizzazione per plausibilità linguistica e sul carattere strutturale delle cosiddette allucinazioni, v. E. BENDER *et al.*, *On the Dangers of Stochastic Parrots. Can Language Models Be Too Big?*, in *Proceedings of the ACM Conference on Fairness, Accountability, and Transparency*, 2021, 610 ss.

distanza tra coerenza interna e verità. L'allineamento semantico resta una procedura di classificazione, non una verifica ontologica. La verità, nel senso epistemico che si vorrebbe rilevante per il diritto, rimane esterna a tali operazioni e dipendente da pratiche di giudizio e di interpretazione[27].

Il rischio che l'*AI Act* non tematizza espressamente diventa così sempre più evidente: la confusione tra coerenza tecnica e verità normativa. La semantica computazionale può garantire precisione formale, stabilità classificatoria o plausibilità statistica, ma nulla di tutto ciò equivale a verità in senso epistemico, né può fungere da fondamento per vincoli giuridici[28]. Ciò che un sistema restituisce come *output* "affidabile" è, al più, un esito coerente rispetto a un modello o a una struttura formale, non un giudizio dotato di validità normativa. Su questa ambiguità si regge una parte significativa della retorica regolatoria europea, che tende a trasfigurare un costrutto tecnico in un presunto fondamento etico.

Questa traslazione, per esempio, emerge con chiarezza nell'art. 10, §3 *AI Act*:

«I set di dati di addestramento, convalida e prova sono pertinenti, sufficientemente rappresentativi e, nella misura del possibile, esenti da errori e completi nell'ottica della finalità prevista. Essi possiedono le

[27] Sui limiti epistemici delle ontologie formali, sulla distinzione tra coerenza interna e verità e sull'esternalità della verità rispetto ai sistemi formali, il rinvio va fatto chiaramente a A. TARSKI, *The Concept of Truth in Formalized Languages*, Oxford University Press, Oxford, 1956, 152 ss.

[28] La distinzione tra plausibilità statistica degli *output* dei sistemi di *machine learning* e verità epistemica, nonché sull'errore categoriale che consiste nel confondere accuratezza predittiva e validità normativa, v. il caso specifico delle raccomandazioni algoritmiche in ambito medico: F. FUNER, *The Deception of Certainty: How Non-Interpretable Machine Learning Outcomes Challenge the Epistemic Authority of Physicians*, in «Journal of Medical Ethics», 2022, 168 ss.; v. anche A. ALDIERI *et al.*, *Consensus Statement on the Credibility of Machine Learning Models*, in «Nature Medicine», 2, 2025, 3 ss.

proprietà statistiche appropriate anche, ove applicabile, per quanto riguarda le persone o i gruppi di persone relativamente ai quali il sistema di IA ad alto rischio è destinato a essere usato».

La disposizione non afferma che dai dati si possa desumere il "vero" in senso epistemico, né che l'accuratezza tecnica equivalga a verità normativa. Però, presuppone implicitamente che la combinazione di rappresentatività, completezza ed "esenzione da errori" sia condizione sufficiente per rendere l'*output* del sistema adeguato a fondare decisioni giuridicamente rilevanti e legittime.

La questione, per altro verso, non può essere risolta evocando l'impiego di strumenti come la *semantic AI*, le ontologie o i modelli linguistici in quanto tali. Si tratta di tecniche che possono risultare funzionali in contesti determinati. Se, però, tali opzioni vengono caricate di una funzione etico-normativa, come se fossero in grado di produrre oggettività o verità idonee a fondare obblighi giuridici o giustificare un processo decisionale automatizzato, allora assistiamo a una commistione tra i piani del discorso. I dati non possiedono un significato intrinseco; le ontologie non garantiscono un accesso privilegiato alla realtà; la semantica computazionale non genera giudizi di verità.

Un'etica regolatoria che assume questi strumenti come fondamento di universalità finisce così per confondere vincoli epistemici e fondamenti normativi[29]. I limiti tecnici e le possibilità operative vengono trattati come se costituissero garanzie di validità prima epistemica e poi etica. Non esiste alcun "dato puro" da cui dedurre principi universali o che consenta di prendere decisioni "giuste": esistono processi sociali, tecnici e politici attraverso cui i dati vengono selezionati, organizzati e interpretati. Insistendo su questa

[29] La tendenza delle strategie regolatorie a scambiare limiti e possibilità tecniche per garanzie di validità etica o giuridica è registrata anche nel *report* OECD, *Regulatory Challenges Raised by Innovation*, Paris, 2021, 13 ss.

illusione, l'*AI Act* rischia di rafforzare una retorica di oggettività che non chiarisce, ma anzi opacizza le scelte politiche e contingenti che sorreggono l'impianto normativo.

4. *Explainability* e semantica artificiale

Uno dei pilastri della regolazione europea dell'intelligenza artificiale, nonché delle differenti proposte etiche nel settore delle tecnologie digitali, è il requisito di *explainability*, inteso come esigenza di trasparenza e di giustificabilità degli *output* prodotti dai sistemi automatizzati. L'*AI Act* assume che tale requisito possa essere soddisfatto mediante strumenti tecnici quali la tracciabilità dei dati, la ricostruzione dei processi decisionali e la formalizzazione delle relazioni semantiche interne ai modelli. La tesi qui sostenuta, ancora una volta, è che questa impostazione rischi di fondarsi su un'ulteriore ambiguità concettuale: l'assimilazione tra tracciabilità causale dei processi computazionali e spiegazione razionale dei contenuti che essi producono. Ne deriva, così, un altro equivoco epistemico, per cui la descrizione tecnica del funzionamento del sistema viene scambiata per una spiegazione nel senso che sembrerebbe rilevante per il diritto.

In primo luogo, occorre provare a distinguere tra causalità e spiegazione. Un sistema di intelligenza artificiale può, almeno in linea di principio, rendere tracciabile la catena causale che conduce a un determinato *output*: l'attivazione di nodi in una rete neurale, la propagazione dei segnali, l'aggiornamento dei pesi, l'influenza di specifici *dataset*. Questa ricostruzione fornisce una descrizione causale rigorosa, (molto lunga, ma) finita e formalmente corretta. Essa non coincide con una spiegazione nel senso umano e, per quanto ci interessa, giuridico del termine.

Innanzitutto, la concatenazione di operazioni elementari che costituisce la causalità computazionale, pur non essendo infinita, è di fatto impraticabile per una mente umana, non solo sul piano cognitivo ma anche su quello funzionale. Percorrerla integralmente richiederebbe tempi, competenze e strumenti incompatibili con qualsiasi uso razionale della spiegazione, se non al prezzo di delegarne l'interpretazione a un altro sistema artificiale, che controlli la concatenazione, e sia a sua volta sottratto alla nostra verifica.

Se ci si chiedesse, ad esempio, perché un LLM abbia prodotto una certa risposta piuttosto che un'altra, la risposta causale consisterebbe in una sequenza di miliardi di operazioni computazionali. Anche ammesso che tale sequenza fosse integralmente accessibile, la sua ricostruzione richiederebbe ben più del tempo biologico di una vita umana e, anche ove fosse possibile, non fornirebbe comunque alcuna spiegazione appagante. Un individuo che avesse dedicato l'intera esistenza a ricostruire la catena causale di un singolo *output* non avrebbe, per ciò stesso, compreso le ragioni di quella "decisione".

Questo dato non mette affatto in discussione la coerenza causale né l'efficacia tecnica dei sistemi di IA. Il problema non è il funzionamento, ma l'inadeguatezza epistemica di questo tipo di tracciabilità rispetto a ciò che, nel linguaggio ordinario e giuridico, si intende per spiegazione. Quando si chiede una spiegazione di una decisione, non si richiede mai l'enumerazione esaustiva delle condizioni causali che l'hanno resa possibile[30]. Nessuno, nel valutare una decisione umana, pretende la ricostruzione della catena

[30] La distinzione tra ricostruzione causale e spiegazione come risposta selettiva e contestuale a *why-questions* è analizzata in B.C. VAN FRAASSEN, *The Scientific Image*, Oxford University Press, New York-Oxford, 1980, 155 ss.; sulla «pragmatics of explanation» v. anche R. CAMPANER, *La causalità tra filosofia e scienza*, Archetipo, Bologna, 2007, 12 ss.

neurobiologica che ha condotto alla produzione di un enunciato. Anche se tale ricostruzione fosse teoricamente disponibile, essa risulterebbe comunque inutilizzabile rispetto alle funzioni che la spiegazione assolve nel contesto del diritto: comprendere, valutare, contestare una decisione. Alla domanda sul perché un giudice abbia deciso in un certo modo o un imputato agito in un certo altro modo, non si risponderebbe mai con l'elenco dei loro stati neurali, né tantomeno fino a una regressione causale potenzialmente infinita.

Questa distinzione emerge con maggiore chiarezza se si adotta la prospettiva della cosiddetta pragmatica della spiegazione. Spiegare non significa elencare tutte le cause, ma selezionare ragioni rilevanti in relazione a una domanda, a un contesto e a uno scopo. La spiegazione opera per sintesi, pertinenza e intelligibilità, non per accumulo causale. In questo senso, una spiegazione causalmente completa sarebbe, paradossalmente, priva di qualsiasi valore esplicativo: troppo lunga, troppo opaca, troppo distante dai criteri di comprensione che fondano la giustificazione razionale. Proprio per questo, essa risulterebbe anche del tutto inutile rispetto alle finalità per cui una spiegazione è richiesta nell'esercizio dei diritti.

Il secondo passaggio dell'argomento riguarda il ruolo della cosiddetta *semantic AI*, di cui si è già detto nel paragrafo precedente. Le tecniche di rappresentazione distribuzionale, come gli *embedding* vettoriali, non catturano il significato delle parole in senso semantico, ma la loro distribuzione statistica all'interno di grandi *corpora* testuali[31]. Il fatto che un modello collochi termini

[31] Sulla "distributional hypothesis" e sul fatto che gli *embedding* codificano regolarità d'uso e co-occorrenze, v. già il testo fondativo che costituisce il possibile sottostante alla maggior parte degli approcci pratici attuali al *Natural Language Processing*, Z. HARRIS, *Distributional Structure*, in «Word», 10, 1954, 146 ss.

come "giustizia" e "diritto" in regioni vicine di uno spazio vettoriale non implica alcuna comprensione concettuale, ma riflette semplicemente una frequente co-occorrenza linguistica. Presentare queste operazioni come "semantiche" contribuisce a rafforzare l'illusione che la trasparenza tecnica equivalga a comprensione, quando si tratta, in realtà, di una forma di sintassi statisticamente raffinata elevata a semantica per convenzione. L'argomento conduce, in sostanza, alla nota obiezione della "stanza cinese" formulata da Searle.

Anche ammettendo una tecnica di *embedding* possa avere un'utilità diagnostica del processo eseguito dalla macchina o comunque descrittiva, ciò non risolve il problema che si è posto nel notare l'opacità delle "decisioni algoritmiche". Rendere visibile una struttura interna del modello non equivale a spiegare perché un certo *output* sia corretto, appropriato o giuridicamente rilevante. La spiegazione richiesta dal diritto non sembra poter consistere in una mappa dello spazio vettoriale o in una formalizzazione ontologica, ma in una giustificazione che consenta di comprendere e ripercorrere le ragioni di una decisione. La confusione tra questi piani alimenta una concezione puramente formale dell'*explainability*, incapace di soddisfare le esigenze di intelligibilità proprie del discorso giuridico[32].

Un terzo passaggio riguarda, infine, proprio le implicazioni giuridiche, in particolare in relazione alla responsabilità e al diritto di difesa nelle decisioni automatizzate. Nel discorso della regolazione digitale si afferma spesso che l'*explainability* sia la

[32] Molte tecniche di XAI forniscono "spiegazioni" descrittive del modello, senza però garantire intelligibilità e giustificazione nel senso richiesto da contesti normativi, v. B. MITTELSTADT, C. RUSSELL, S. WATCHER, *Explaining Explanations in AI*, in *Proceedings of the ACM Conference on Fairness, Accountability, and Transparency*, 2019, 279 ss

condizione necessaria per rendere le decisioni compatibili con i diritti fondamentali. Tuttavia, se la spiegazione fornita si riducesse a una ricostruzione tecnica del processo computazionale, essa non consentirebbe né di attribuire responsabilità né di esercitare un diritto di difesa in senso proprio (in ipotesi, contro una "decisione" avversa a un diritto). Difendersi da una decisione significa poter comprendere le ragioni che l'hanno giustificata, valutarne la correttezza e contestarne i presupposti normativi o fattuali. Una catena causale opaca, per quanto formalmente tracciabile, non assolve nessuna di queste funzioni. Impugnare una decisione fondata su miliardi di operazioni computazionali, "spiegata" attraverso il peso di sequenze simboliche o vettoriali, solleva inevitabilmente la domanda: quale utilità effettiva ha una simile spiegazione ai fini dell'esercizio del diritto di difesa?

L'*explainability* tecnica rischia così di funzionare come un dispositivo di rassicurazione simbolica. Essa produce un'impressione di controllo e di trasparenza senza restituire al soggetto destinatario una reale possibilità di comprendere e contestare la decisione. In termini foucaultiani, volendo ancora accogliere l'impostazione che sostiene le ragioni della normativa digitale, la regolazione stessa costruisce un dispositivo che governa attraverso la visibilità dei processi, senza garantire l'intelligibilità del loro senso. La trasparenza diventa allora un fine in sé, sganciato dalla funzione giustificativa che dovrebbe assolvere.

L'*explainability* richiesta dal diritto europeo tende a operare come un atto meramente performativo: costruisce la percezione di responsabilità e controllo, senza fornire una spiegazione nel senso pragmatico e giuridicamente rilevante del termine. Insistendo sulla spiegabilità come requisito tecnico, l'*AI Act* rischia

di promettere una forma di comprensione che la tecnologia non è in grado di offrire, ampliando la distanza tra la retorica regolatoria sulle decisioni automatizzate e le condizioni effettive della responsabilità e del diritto di difesa[33].

Ciò non implica una critica all'uso dell'IA nei processi decisionali in quanto tali. Al contrario, significa legittimarli senza pretendere che soddisfino requisiti che non possono soddisfare, recuperando – ove necessario – garanzie propriamente umane in una fase successiva: nel riesame, nell'appello, nella valutazione umana *ex post* dei casi controversi.

5. L'aporia epistemologica dell'*ethics by design*

Tra le ambizioni più problematiche dell'*AI Act* vi è, come anticipato, l'idea che i valori etici e i principi giuridici possano essere incorporati direttamente nella progettazione tecnica dei sistemi di intelligenza artificiale[34]. Non si tratta soltanto di una proposta ope-

[33] Sul rapporto tra decisioni automatizzate, "meaningful information" e limiti del "diritto alla spiegazione" nel GDPR, v. S. WACHTER, B. MITTELSTADT, L. FLORIDI, *Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation*, in «International Data Privacy Law», 7, 2017, 76 ss.

[34] L'idea di incorporare valori etici nella progettazione tecnica dei sistemi di IA e sul suo ruolo nell'*AI Act*, sin dalla comunicazione della COMMISSIONE EUROPEA, *Libro bianco sull'intelligenza artificiale - Un approccio europeo all'eccellenza e alla fiducia*, COM(2020) 65 final, e *Promuovere un approccio europeo all'intelligenza artificiale*, COM(2021) 205 final; cfr. T. EVAS (a cura di), *European Framework on Ethical Aspects of Artificial Intelligence, Robotics and Related Technologies*, EPRS, Brussels. 2020: «the common EU regulatory action on AI ethics could be the 'so-called silver bullet in the EU's strategy to 'catch up' with the USA or China» (18). Quanto all'*AI Act*, per i sistemi ad alto rischio, l'art. 9 impone un *risk management by design* orientato alla prevenzione dei rischi per i diritti fondamentali; l'art. 10 richiede *fairness and non-discrimination by design*

rativa, ma di una tesi implicita di natura epistemologica: che un contenuto determinato di valori complessi – come equità, dignità, proporzionalità o trasparenza – possa essere catturato, fissato e reso operabile all'interno di specifiche tecniche di un *software*.

L'*ethics by design* si presenta, in primo luogo, come una reazione al riduzionismo della tecnologia: essa afferma di voler preservare il senso autentico dei valori, evitando che essi si perdano in un modo dominato da un disumano meccanicismo. Però, nel momento stesso in cui quest'approccio tenta di "incorporarli" nei sistemi tecnologici, è costretto a trattarli come entità definibili, formalizzabili e stabilizzabili. Ciò che viene così presupposto è che esista un contenuto normativo determinato del valore, identificabile *ex ante* e traducibile in regole operative. Ma è proprio questa assunzione a entrare in tensione con la "natura" dei valori che si vorrebbero salvaguardare.

Valori come equità o dignità non funzionano come istruzioni eseguibili, né come proprietà intrinseche di uno stato di cose. Essi sembrerebbero operare come criteri aperti di giudizio, la cui funzione è orientare la valutazione di situazioni concrete, spesso conflittuali, attraverso pratiche argomentative e interpretative. Il loro contenuto non è mai dato una volta per tutte, ma emerge nel confronto tra casi, contesti e interessi divergenti. Pretendere di fissarne il significato "vero" in una fase di progettazione di una piattaforma digitale significa, quindi, trasformarli in qualcosa di ontolo-

attraverso il controllo e la mitigazione dei *bias* nei *dataset*; l'art. 13 prescrive *transparency and explainability by design*; l'art. 14 introduce una *human oversight by design* mediante strumenti tecnici di controllo umano significativo. Il considerando 27 richiama i sette principi della *Trustworthy AI*, che operano come quadro etico-interpretativo degli obblighi vincolanti. Per i sistemi non ad alto rischio, l'art. 95 promuove codici di condotta volontari che estendono l'approccio *ethics by design* a inclusività, sostenibilità e accessibilità.

gicamente diverso da quel che si professa che essi siano: da criteri di giustificazione essi diventano vincoli tecnici.

Questa trasformazione non è un effetto collaterale evitabile, ma una condizione strutturale dell'operazione che si vuole compiere. Per diventare operabili, i valori devono essere tradotti in parametri, soglie, metriche o funzioni obiettivo. Ma, così facendo, essi cessano di operare come valori in senso etico o normativo e diventano elementi di una procedura di ottimizzazione. In tal modo, l'etica incorporata non è più il luogo del giudizio, bensì il risultato di una scelta progettuale che seleziona una particolare interpretazione del valore e la cristallizza nel sistema. Se si voleva preservare l'essere umano dalla tecnica, si è tecnicizzata, in modo improprio e opaco, una sfera del tutto umana.

Il caso dell'equità, da questo punto di vista, può risultare emblematico. Nelle scienze dure l'equità viene spesso tradotta in metriche computabili – come la *statistical parity* o altre nozioni di *fairness* – che consentono di misurare e ottimizzare determinati comportamenti del modello. Questo passaggio, tuttavia, non risolve il problema del significato dell'equità, ma lo aggira. L'ottimizzazione di una metrica garantisce, al più, la conformità a una specifica formalizzazione del valore, lasciando sullo sfondo – e spesso rendendo invisibili – le alternative concettuali e i conflitti normativi sottostanti. Tali metriche sono molto utili, ma sarebbe fallace la pretesa che esse possano esaurire o incarnare l'equità in senso sostanziale, riducendo una nozione intrinsecamente controversa a una serie di vincoli tecnici[35].

[35] I limiti delle metriche di equità mostrano una impossibilità di identificare una nozione computazionale univoca di *fairness*: v. P. KIM, M. RAGHAVAN, *Limitations of the "Four-Fifths Rule" and Statistical Parity*, in «Georgetown Law Technology Review», 8, 2024, 93 ss., dove l'equità viene ridotta a un vincolo formale computabile. Per esempio, il progetto *Fairlearn* mostra come l'equità venga tradotta in metriche computabili e vincoli di ottimizzazione all'interno dei sistemi di

La contraddizione dell'*ethics by design*, quindi, emerge sempre più nitidamente: più essa pretende di rispettare l'irriducibilità dei valori umani, più è costretta a ridurli in elementi scomposti per renderli tecnologicamente trattabili. Il *by design* non fallisce perché semplifica – ogni tecnica, per fortuna, semplifica –, ma perché pretenderebbe di evitare la semplificazione (difendendo l'irriducibilità dell'umano) proprio nel momento in cui la rende necessaria e definitiva (trasformando i valori umani in meri parametri). Ciò che viene presentato come il "senso autentico" del valore è, in realtà, il risultato di una scelta contingente, elevata a fondamento normativo attraverso la mediazione tecnica.

Questa aporia non si dissolve neppure se si accoglie l'idea, di matrice teorico-generale, che i valori funzionino come precetti di ottimizzazione. Anche in questa prospettiva, resta aperto il problema di chi determini il contenuto del precetto, di come si identifichi lo spazio dell'ottimo nella ponderazione e di come si risolvano i conflitti tra valori incompatibili. L'ottimizzazione presuppone una funzione obiettivo già definita, ma è proprio la definizione di tale funzione a essere oggetto di decisione normativa e politica, non di scoperta tecnica.

L'*ethics by design* non rappresenta tanto una soluzione avanzata al problema della regolazione etica dell'intelligenza artificiale, quanto il sintomo di una tensione irrisolta. Essa tenta di sottrarre i valori alla loro dimensione conflittuale e interpretativa, per collocarli in uno spazio di operatività tecnica che promette certezza e neutralità. Ma, per questa via, finisce per svuotare i valori della

machine learning: pur riconoscendo il carattere non univoco della *fairness*, l'approccio finisce per ridurre il valore a conformità tecnica a specifiche formalizzazioni, senza esaurirne il significato normativo sostanziale (H. WEERTS *et al.*, *Fairlearn: Assessing and Improving Fairness of AI Systems*, in «Journal of Machine Learning Research», 24, 2023, 1–8).

funzione che li renderebbe tali: quella di orientare il giudizio umano in situazioni in cui nessuna soluzione può essere garantita *ex ante* come quella "vera" o quella "giusta".

6. La "natura" *technology-driven* dei valori e le sue implicazioni

Assumere che i valori siano, almeno in parte, *technology-driven* significa riconoscere che essi non preesistono alle configurazioni tecniche come criteri stabili e autosufficienti, ma emergono, si trasformano e acquistano rilevanza attraverso le pratiche sociali che le tecnologie rendono possibili. In questa prospettiva, il diritto non può illudersi di fissare i valori *ex ante* in un codice, né di determinarne una volta per tutte il contenuto[36]. Ogni tentativo di incorporare valori presentati come universali all'interno dei sistemi tecnologici presuppone una stabilità che non esiste. Nozioni come dignità, trasparenza o equità non precedono l'esperienza tecnica come parametri immutabili, ma vengono continuamente riplasmate dagli usi, dai conflitti e dalle mediazioni che le tecnologie producono.

Questo punto impone un chiarimento ulteriore, spesso trascurato: non solo i valori non sono dati *a priori*, ma neppure il diritto possiede una funzione naturale o necessaria. L'idea che il diritto *serva* sempre a realizzare determinati valori, a orientare la tecnologia o a garantire diritti fondamentali o a garantire la pace sociale è essa stessa una costruzione teorica storicamente situata. Il diritto è

[36] Sulla co-produzione di valori, pratiche sociali e configurazioni tecniche, v. T. PINCH, W. BIJKER, *The Social Construction of Facts and Artefacts: or How the Sociology of Science and the Sociology of Technology might Benefit Each Other*, in «Social Studies of Science», 14(3), 1984, 399 ss.; nonché diffusamente in L. WINNER, *Do Artifacts Have Politics?*, cit.

uno strumento tra altri strumenti di regolazione dei conflitti sociali; non interviene, però, su tutti i conflitti, ma soltanto su quelli che una determinata comunità politica decide di giuridificare, cioè di tradurre in regole, procedure e decisioni istituzionali. I suoi scopi e le sue diverse efficacie dipendono da scelte politiche e da assetti di potere contingenti, non da una vocazione intrinseca.

Allora, il possibile errore del paradigma regolatorio europeo non consiste soltanto nel ritenere che i valori possano essere tradotti in specifiche tecniche senza perdere il loro significato, ma anche nel sovrainvestire il diritto di una funzione che esso non può garantire, o almeno non sempre. L'*AI Act* sembra muovere dall'aspettativa che la regolazione giuridica debba necessariamente governare lo sviluppo tecnologico e correggerne gli esiti eticamente problematici. Questa aspettativa trascura però un dato elementare: il diritto non sempre governa la tecnologia in quanto tale, ma solo alcuni degli effetti e dei conflitti che essa produce, e lo fa con modalità e secondo fini che restano sempre contestabili[37].

La tesi della natura *technology-driven* dei valori si fonda su un presupposto ormai acquisito nelle scienze sociali e nella filosofia della tecnica, sia pur nei differenti approcci che si possono distinguere pur nella condivisione di un assunto comune: la tecnologia non è neutrale. Gli artefatti incorporano assetti di potere, strutturano possibilità d'azione, orientano percezioni e pratiche. I valori, in questo senso, non si collocano "fuori" dalla tecnologia, in alto, in attesa di essere implementati, ma vengono co-prodotti dalle configurazioni tecniche e dai contesti sociali in cui esse operano. Non è la tecnica a ricevere passivamente i valori; è il contesto tecnologico a contribuire a ridefinirne il significato dei valori.

[37] Quanto ai limiti strutturali della regolazione giuridica rispetto allo sviluppo tecnologico, v. M. HILDEBRANDT, *Law for Computer Scientists and Other Folk*, Oxford University Press, 2020, 9 ss. sulle «frontiers of law in an onlife world».

Ne consegue che il diritto non può essere concepito come un ingegnere morale incaricato di progettare in anticipo l'assetto valoriale dei sistemi tecnologici. Può, al massimo, essere impiegato come uno degli strumenti attraverso cui una società decide di intervenire su determinati conflitti generati dall'uso delle tecnologie. Questo intervento non risponde a una funzione eticamente necessaria, ma a obiettivi scelti dal decisore politico: proteggere alcune posizioni soggettive, ridurre specifici rischi, distribuire responsabilità, preservare equilibri istituzionali. Stabilire quali obiettivi perseguire, e se il diritto sia lo strumento più adeguato a farlo, non è una questione giuridica in senso stretto, ma una decisione eminentemente politica.

In questa prospettiva, anche modelli come le *regulatory sandboxes* o le forme di *governance* adattiva non rappresentano soluzioni intrinsecamente superiori, ma opzioni strumentali[38]. Possono risultare utili in determinati contesti, quando si ritiene opportuno osservare e correggere *ex post* gli effetti delle tecnologie, ma non costituiscono un modello normativo necessario né una risposta definitiva al rapporto tra diritto, tecnologia e valori.

Da un punto di vista teorico più generale, ciò richiede l'abbandono di una concezione astrattamente finalistica del diritto. Il diritto non possiede (sempre e solo) una funzione naturale di garanzia dei valori o di governo della tecnica. Può essere utilizzato per intervenire su specifici conflitti, traducendoli in categorie giuridiche e sottoponendoli a procedure decisionali, ma resta sempre un dispositivo parziale, selettivo e storicamente situato. Esso non si

[38] La natura contingente e non necessaria delle *regulatory sandboxes* nell'abito regolatorio in questione è oggetto dell'analisi di M. ESPAÑOL *et al.*, *Regulatory Sandboxes for AI*, in «Cambridge Forum on AI Law and Governance», 2025, 6 ss.; cfr. anche il *report* della Banca Mondiale, I. JENÍK, S. DUFF (a cura di), *How to Build a Regulatory Sandbox*, Washington DC, 2020, 4 ss.

limita a inverare o riflettere i valori di una società, ma contribuisce a produrli, selezionando ciò che diventa giuridicamente rilevante e ciò che resta fuori dal diritto stesso.

Riconoscere il carattere (anche) *technology-driven* dei valori e la "natura" strumentale del diritto non implica rinunciare alla regolazione dell'intelligenza artificiale. Implica, piuttosto, rinunciare all'illusione che il diritto possa fornire risposte necessarie o definitive a tutti problemi posti dalla tecnologia. La regolazione appare allora per ciò che è: una pratica contingente, attraverso cui una comunità decide quali conflitti affrontare ricorrendo al diritto e quali obiettivi perseguire, nella consapevolezza che né i valori né il diritto possiedono una stabilità garantita *ex ante*.

7. Conclusioni

L'analisi dell'*AI Act* consente di chiarire l'equivoco che ne attraversa l'intero impianto: l'illusione dell'etica regolatoria. L'idea che il diritto possa imporre alla tecnologia valori universali e intrinseci presuppone che tali valori siano dati *ex ante*, semanticamente stabili e tecnicamente traducibili. Questa assunzione, diffusa nel discorso politico e in una parte della dottrina, poggia su basi concettuali molto fragili.

I dati non costituiscono informazione semantica fattuale, ma acquistano significato solo all'interno di pratiche interpretative e contesti sociali determinati. L'*explainability* tecnica offre tracciabilità causale, peraltro utilissima in certi casi, ma non offre spiegazioni giuridicamente sensate: la ricostruzione dei processi computazionali non produce ragioni comprensibili, né giustificazioni contestabili. I valori, infine, non sopravvivono alla loro riduzione a parametri computabili senza subire una trasformazione sostanziale.

Ciò che nasce come criterio di giudizio e di conflitto viene convertito in una condizione tecnica di ottimizzazione. L'etica *by design* produce così una versione tecnicizzata e impoverita dell'etica, accompagnata da un irrigidimento procedurale e burocratico che grava sull'innovazione e neutralizza la complessità che vorrebbe governare.

Questa impostazione genera prevalentemente effetti simbolici. L'etica regolatoria sembra operare come una semantica rassicurante: segnala intenzioni, promette controllo, costruisce legittimazione delle istituzioni. L'*AI Act* assume, allora, la forma di un dispositivo politico che mira a compensare una debolezza strutturale dell'Unione europea nel campo dell'innovazione tecnologica. A differenza di quanto si è talvolta sostenuto con riferimento al GDPR, nel campo dell'intelligenza artificiale l'"effetto Bruxelles" appare soprattutto come un obiettivo dichiarato, la cui capacità di produrre reali convergenze normative resta incerta.

Il punto di approdo di questo contributo non consiste nell'individuare un diverso modello regolatorio o nel recuperare o scoprire una funzione più nobile del diritto che altri avrebbero trascurato. Tuttavia, si propone un mutamento di prospettiva. I valori non precedono la tecnologia come criteri esterni pronti per l'implementazione, ma vengono continuamente riplasmati anche dalle configurazioni tecnologiche e dalle pratiche sociali che esse rendono possibili. I valori risultano, almeno in parte, *technology-driven*. La relazione tra tecnica e valori assume la forma di una coproduzione, incompatibile con l'idea di un'etica preconfezionata che si inietta nella progettazione ingegneristica.

Questo quadro impone anche una revisione di un'immagine del diritto molto ancorata a impostazioni tradizionali e, forse, superate. Il diritto non possiede una funzione naturale, né una vocazione intrinseca a garantire valori, orientare la tecnologia o realiz-

zare la giustizia. Opera come strumento selettivo di regolazione dei conflitti, intervenendo solo su quelli che una comunità politica decide di giuridificare. La scelta dei conflitti rilevanti, delle regole applicabili e degli obiettivi perseguiti dipende ora da decisioni politiche contingenti, ora semplicemente dal caso. L'etica regolatoria europea sovrastima tanto la traducibilità tecnica dei valori quanto la capacità del diritto di sostenere il carico funzionale che gli viene attribuito.

La questione più rilevante è stabilire *quando*, *come* e *perché* ricorrere al diritto per regolare determinati conflitti che le tecnologie ridefiniscono o generano *ex novo*. Rinunciare all'illusione dell'etica incorporata consente di sottrarre la regolazione a una pretesa di necessità che la rende fragile e ideologica. Ogni intervento normativo appare allora per ciò che è: una decisione situata, reversibile e contestabile.

Smascherare l'illusione dell'etica regolatoria è anche un modo per restituire sobrietà al discorso giuridico. D'altronde, ridimensionare le pretese del diritto consente di chiarirne i limiti e, proprio per questo, le possibilità effettive.

Tecniche di normazione dell'intelligenza artificiale tra hard e soft law: una lettura nichilista

di *Nicola Colacino*[*]

SOMMARIO: 1. Intelligenza artificiale, trasformazione della società e ruolo del diritto. – 2. Normare l'intelligenza artificiale tra pretese di autonomia e derive nichiliste. – 3. I primi tentativi di regolamentazione internazionale e l'assenza di un approccio globale condiviso. – 4. La Convenzione quadro del Consiglio d'Europa sull'intelligenza artificiale, i diritti umani, la democrazia e lo Stato di diritto tra ambizioni e lacune. – 5. Considerazioni conclusive: i modelli di regolamentazione europei come argine al primato della tecnica.

1. Intelligenza artificiale, trasformazione della società e ruolo del diritto

Il sempre più frequente e pervasivo ricorso ai sistemi di intelligenza artificiale (d'ora in avanti, anche IA) in numerosi ambiti dell'attività umana ha innescato una profonda trasformazione della società, contribuendo a ridefinire talune nozioni e categorie concettuali con evidenti implicazioni sotto il profilo etico, filosofico e giuridico[1]. Benché sia incontestata l'utilità di tali sistemi per la

[*] Professore ordinario di Diritto internazionale, CASD - Scuola Superiore Universitaria.

[1] I contributi della dottrina giuridica italiana e internazionale in materia sono già numerosi e riflettono l'ampiezza del dibattito corrente, incentrato sugli aspetti definitori e sugli elementi fondativi della disciplina. Si vedano, tra gli altri, G. FINOCCHIARO, *Diritto dell'intelligenza artificiale*, Zanichelli, 2024; R. RAZZANTE (a cura di), *Manuale sull'Intelligenza Artificiale*, Giappichelli, Torino, 2024; B. CU-

semplificazione e l'efficientamento dei processi produttivi e, in termini generali, per il raggiungimento di un più elevato livello di

STERS, E. FOSCH-VILLARONGA (*eds.*), *Law and Artificial Intelligence: Regulating AI and Applying AI in Legal Practice*, T.M.C. Asser Press, The Hague, 2022; J. LEE, *Artificial Intelligence and International Law*, Springer, Singapore, 2022; C. KERRIGAN (*ed.*), *Artificial Intelligence: Law and Regulation*, Edward Elgar, Cheltenham (UK)–Northampton (MA), 2022; A. PAJNO, F. DONATI, A. PERRUCCI (a cura di), *Intelligenza artificiale e diritto: una rivoluzione?*, 2 voll., Il Mulino, Bologna, 2022; G. SARTOR, *L'intelligenza artificiale e il diritto*, Giappichelli, Torino, 2022; G. ALPA, *L'intelligenza artificiale: Il contesto giuridico*, Mucchi, Modena, 2021; S. CHESTERMAN, *We, The Robots?: Regulating Artificial Intelligence And The Limits Of The Law*, CUP, Cambridge, 2021; S. J. THOMPSON, *Machine Law, Ethics, and Morality in the Age of Artificial Intelligence*, IGI Global, Hershley PA, 2021; J. DE BRUYNE; C. VANLEENHOVE, *Artificial Intelligence And The Law*, Intersentia, Cambridge, 2021; S. DEAKIN, C. MARKOU (*eds.*) *Is Law Computable?: Critical Perspectives on Law and Artificial Intelligence*, Hart Publishing, Oxford, 2020; K. D. ASHLEY, *Artificial Intelligence and Legal Analytics: New Tools for Law Practice in the Digital Age*, CUP, Cambridge, 2017. Per una disamina delle principali questioni legate all'impatto etico e alle correlate esigenze di regolamentazione giuridica dei sistemi di IA si rinvia a L. A. DIMATTEO, C. PONCIBÒ, M. CANNARSA (*eds.*), *Handbook of Artificial Intelligence: Global Perspectives on Law and Ethics*, CUP, Cambridge, 2022. Dalla lettura dei diversi contributi che compongono il volume emerge la difficoltà di adattare ai sistemi di IA taluni approcci teorici e categorie giuridiche tradizionali, in ragione sia della novità del fenomeno, sia della sua tendenziale resistenza alla codificazione e del costante progresso tecnologico che lo caratterizza. Al di là del problema dell'adeguatezza degli strumenti di regolazione, giocoforza in bilico tra la segnalata opera di adattamento dei principi e delle regole vigenti e l'elaborazione di nuovi istituti e costrutti teorici, nel volume trova spazio anche una riflessione generale sul ruolo del giurista, che prende in considerazione il rapporto tra la conoscenza e l'applicazione delle norme giuridiche e l'impatto dell'evoluzione tecnologica: "[w]hen lawyers had no interest in technology, thinking like a lawyer was inward-looking, doctrinal and guided by general principles. By the time lawyers saw technology as a challenge, thinking like a lawyer needed to become more 'regulatory', more policy-oriented, and more outward-looking (with a view to learning from economics, sociology and philosophy)" (R. BROWNSWORD, *Foreword*, ivi, *xxiii*). Proprio dalle trasformazioni della funzione normativa e del ruolo del giurista di fronte all'IA traggono spunto le considerazioni che animano il presente lavoro.

benessere economico, il loro impiego ormai capillare è considerato un fattore di rischio per la tutela dei valori democratici e l'uguaglianza sociale[2].

Sennonché, se sul piano etico e filosofico gli interrogativi posti dall'avvento dell'IA e, in particolare, dai nuovi e sofisticati modelli di natura generativa, appaiono per lo più riducibili al dilemma morale – già emerso nel dibattito sulla clonazione[3] – dell'ammissibilità della creazione di un soggetto in tutto simile all'essere umano, in grado di simularne il pensiero, e quindi astrattamente suscettibile del medesimo livello di considerazione[4], su quello giuridico il

[2] I considerando (4) e (5) del regolamento (UE) 2024/1689 del Parlamento europeo e del Consiglio del 13 giugno 2024 sull'intelligenza artificiale, più noto come "AI Act", chiariscono che "l'IA consiste in una famiglia di tecnologie in rapida evoluzione che contribuisce al conseguimento di un'ampia gamma di benefici a livello economico, ambientale e sociale nell'intero spettro delle attività industriali e sociali […]", ma che può "nel contempo, a seconda delle circostanze relative alla sua applicazione, al suo utilizzo e al suo livello di sviluppo tecnologico specifici, comportare rischi e pregiudicare gli interessi pubblici e i diritti fondamentali tutelati dal diritto dell'Unione. Tale pregiudizio può essere sia materiale sia immateriale, compreso il pregiudizio fisico, psicologico, sociale o economico". Sul punto, ammonisce K. GEORGIEVA, *AI Will Transform the Global Economy. Let's Make Sure It Benefits Humanity*, in «IMF Blog», 14 gennaio 2024, https://www.imf.org (data di accesso 20 marzo 2025): "[i]n most scenarios, AI will likely worsen overall inequality, a troubling trend that policymakers must proactively address to prevent the technology from further stoking social tensions. It is crucial for countries to establish comprehensive social safety nets and offer retraining programs for vulnerable workers. In doing so, we can make the AI transition more inclusive, protecting livelihoods and curbing inequality".

[3] L. R. KASS, J. K. WILSON, *The Ethics of Human Cloning*, AEI Press, Washington, 1998.

[4] C. NOVELLI, G., BONGIOVANNI, G. SARTOR, *A conceptual framework for legal personality and its application to AI*, in «Jurisprudence», 2021, 194-219; J. DOOMEN, *The artificial intelligence entity as a legal person*, in *Information & Communications Technology Law*, 2023, 277-287. V. anche V. A. J. KURKI, *The Legal Personhood of Artificial Intelligences*, in *A Theory of Legal Personhood*, Oxford Legal Philosophy, Oxford, 2019, 175-190, secondo cui "if we assume that some AIs are of ultimate value [nella misura in cui esse "achieve a status of moral considerability"], then they can

ventaglio delle questioni rilevanti appare più ampio ed eterogeneo.

Basti pensare alla varietà di conseguenze che l'introduzione dei sistemi di IA è in grado di determinare sullo statuto giuridico dell'individuo-lavoratore o dell'individuo-consumatore, a tacere della progressiva, ma irreversibile, modificazione del rapporto tra cittadino e poteri pubblici – da quello giudiziario a quello amministrativo di controllo in materia fiscale, solo per citarne alcuni.

Si tratta di trasformazioni sociali ancora non chiaramente percepibili nella loro complessità – con le quali il giurista fatica, quindi, a misurarsi – e dall'esito non prevedibile, data l'impossibilità di stabilire in anticipo sia il ritmo di sviluppo delle future applicazioni dell'IA, sia il loro grado di incidenza progressiva, apprezzabile in funzione della maggiore o minore attitudine della società ad adattarvisi senza rinnegare i propri valori o mettere in discussione il proprio assetto istituzionale.

Il precipitato giuridico delle numerose possibilità di impiego dell'intelligenza artificiale compone, pertanto, un quadro ancora frammentario. Se si eccettua il tentativo di regolamentazione promosso dall'Unione europea con l'AI Act[5] e il suo approccio alla materia basato sul rischio[6], non è ancora rinvenibile un chiaro criterio ispiratore delle misure finora proposte.

hold claim-rights; we can owe duties to them, and our duties do not merely pertain to them. We can therefore conclude that they can be passive legal persons" (178).

[5] V. *supra*, nota 2. Si tornerà più diffusamente sulle caratteristiche e le finalità di tale atto legislativo *infra*, par. 5. Per un inquadramento generale sui relativi contenuti si rinvia a N. Th. Nikolinakos, *EU Policy and Legal Framework for Artificial Intelligence, Robotics and Related Technologies - The AI Act*, Springer, Cham, 2023.

[6] E. Cirone, *L'*AI Act *e l'obiettivo (mancato?) di promuovere uno standard globale per la tutela dei diritti fondamentali*, in «Quaderni AISDUE», f.s. 2, 2024, *La nuova disciplina UE sull'intelligenza artificiale*, 1-20, in particolare 6, dove sottolinea come il regolamento sull'intelligenza artificiale dell'UE rappresenta "uno strumento che, pur favorendo l'integrazione nell'ambito del mercato unico, adot-

Parimenti, a livello internazionale, non si registra una sufficiente convergenza sulla scelta degli strumenti e delle procedure da seguire. Sono state avanzate proposte sia di hard law, sia di soft law, la cui efficacia, tuttavia, appare ancora incerta: ciò è imputabile, nel primo caso, sia alla segnalata rapidità di evoluzione dei sistemi di IA (e al ritardo che fatalmente caratterizzerebbe l'avvio di procedimenti di regolamentazione formali), sia all'eterogeneità e all'ampia diffusione delle rispettive applicazioni – fattori che contrastano con la limitatezza dell'ambito di operatività degli strumenti disponibili; nel secondo caso, all'intrinseca debolezza degli atti di soft law e alla loro inidoneità a imporre l'osservanza generalizzata di determinati standard di condotta in assenza di una volontà conforme. Un'ipotesi alternativa è rappresentata dal ricorso a una normazione per principi[7], comunque non esente dal rischio di manipolazioni o divergenze interpretative.

Considerata la quantità di settori in cui i sistemi di IA appaiono destinati a venire in rilievo, è probabile, peraltro, che le forme e gli strumenti di regolamentazione assumano con il tempo una con-

ta un approccio specifico (*rectius*, peculiare): garantisce certezza giuridica traendo ispirazione dalla regolamentazione della sicurezza dei prodotti – beneficiando di decenni di esperienza acquisita con tali strumenti che sono diventati uno standard globale – e, al contempo, inserisce in tale cornice giuridica la tutela dei diritti fondamentali come requisito essenziale di *compliance*".

[7] In base alla ricognizione effettuata dal Berkman Klein Center for Internet & Society nel 2020, la disciplina giuridica dell'IA potrebbe essere radicata in otto principi di base, segnatamente: tutela della privacy, responsabilità, sicurezza e protezione, trasparenza e intelligibilità, equità e non discriminazione, controllo umano della tecnologia, responsabilità professionale e rispetto dei valori umani. J. FJELD, N. ACHTEN, H. HILLIGOSS, A. NAGY, M. SRIKUMAR, *Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-based Approaches to Principles for AI*, Berkman Klein Center for Internet & Society, Cambridge, 2020, https://dash.harvard.edu/server/api/core/bitstreams/c8d686a8-49e8-4128-969c-cb4a5f2ee145/content.

figurazione sempre più specifica, variando da settore a settore in funzione dell'individuazione del punto di equilibrio tra interessi economici e garanzie di tutela sociale. Al netto di tale prospettiva, non mancano i tentativi di ricognizione e classificazione degli ambiti tematici su cui la disciplina giuridica dell'intelligenza artificiale tende attualmente a concentrarsi. Si tratta delle questioni legate alla governance dei sistemi di IA, di quelle in materia di privacy e di sicurezza e, infine, dei profili di responsabilità e di tutela dell'utenza.

Non pare, invece, suscitare il medesimo interesse il problema dell'attribuzione della personalità giuridica all'IA, che purtuttavia presenta un rilievo specifico ove si intenda ragionare sull'autonomia dei processi decisionali non umani e sui profili di responsabilità per gli effetti potenzialmente pregiudizievoli collegati agli esiti di tali processi. In proposito, è stato correttamente messo in luce come non esistano elementi ostativi formali al riconoscimento della personalità giuridica in capo a specifiche applicazioni dell'IA, benché ciò possa ritenersi ancora prematuro, data l'assenza di un livello di autocoscienza sufficiente per poter valutare dette applicazioni alla stregua di centri autonomi di imputazione giuridica[8].

Tale condivisibile approccio è certamente debitore del contributo kelseniano alla concezione della persona giuridica, unica forma realmente valutabile dalla scienza del diritto ai fini dell'attribuzione della soggettività e della titolarità di situazioni attive e passive: "[l]'unità di doveri e di diritti soggettivi, cioè l'unità delle norme giuridiche qui considerate, le quali formano una persona

[8] Nell'opinione di J. DOOMEN, *The artificial intelligence entity as a legal person*, cit., 283, "[a] quest for the 'right' definition (or 'essence') of 'personhood' may be in vain". Ciò in quanto "there is no need to refer to one or more absolute criteria, the question to which entities legal personhoodis bestowed being decided by the legislature".

fisica, è data dal fatto che è il comportamento del medesimo uomo a formare il contenuto di questi diritti e doveri, che è cioè il comportamento dello stesso uomo ad essere determinato da queste norme giuridiche. La cosiddetta persona fisica non è quindi un uomo, bensì l'unità personificata delle norme giuridiche che attribuiscono doveri e diritti al medesimo uomo. Non è una realtà naturale, bensì una costruzione giuridica creata dalla scienza del diritto, un concetto ausiliario nella descrizione di fattispecie giuridicamente rilevanti. In questo senso la cosiddetta persona fisica è una persona giuridica"[9].

In questa prospettiva, sembra corretto ragionare in concreto sulla possibilità di attribuire la personalità giuridica alle applicazioni dell'intelligenza artificiale "forte" o "generale", vale a dire quelle che, raggiungendo specifiche soglie di pensiero indipendente e una sufficiente capacità di analisi e decisione, "can in relevant respects act like a human being"[10]. La maggior parte delle tecnologie di IA attualmente in uso non soddisfa ancora queste caratteristiche. Peraltro, quando si discute di attribuzione della personalità giuridica, occorre tener conto che anche le applicazioni di IA forte non possono essere considerate alla stregua di entità completamente indipendenti dal punto di vista giuridico. Non è ancora ipotizzabile, infatti, che un sistema di IA possa costituire un membro del corpo sociale, suscettibile di essere punito o educato per le condotte ad esso imputabili. Il riconoscimento della titolarità di diritti e doveri, in questa fase, deve piuttosto essere ricondotto a modelli già utilizzati per altri enti dotati di personalità giuridica, ancorché tale equiparazione non possa essere ritenuta pienamente soddisfacente, dal momento che i continui progressi tecnologici tendono

[9] H. KELSEN, *La dottrina pura del diritto* (a cura di M. G. LOSANO), Einaudi, Torino, 2021, 198.

[10] V. A. J. KURKI, *The Legal Personhood of Artificial Intelligences*, cit., 177.

ad avvicinare le applicazioni di IA più avanzate ai tratti degli esseri umani (come il pensiero e la capacità di decisione). Sembra utile, pertanto, provare a delineare schemi differenti di personalità giuridica in base alle caratteristiche, modalità di impiego e finalità dei vari sistemi di IA.

2. Normare l'intelligenza artificiale tra pretese di autonomia e derive nichiliste

La difficoltà di prevedere con ragionevole approssimazione il ritmo di evoluzione tecnologica dei sistemi di IA e il loro futuro impatto sull'organizzazione della società è l'incognita con cui qualsiasi tentativo di normazione è destinato a confrontarsi. L'assenza di una prospettiva di applicazione durevole e la mutazione della funzione della norma regolatrice a mero compendio transitorio evocano lo spettro del nichilismo giuridico, espressione utilizzata da Natalino Irti per descrivere la progressiva riduzione della complessità del fenomeno giuridico a mera attività di produzione di regole mediante procedure "indifferenti verso ogni contenuto"[11], ma parimenti adatta a rappresentare la crisi di senso che il giurista

[11] N. IRTI, *Nichilismo giuridico*, Laterza, Bari, 2004, *vi*. Irti individua nell'esasperazione del formalismo e nel costante incremento quantitativo dei rapporti i tratti distintivi del nichilismo giuridico. Le norme, espropriate del loro fondamento etico sostanziale, sono ridotte a mere formule per la gestione di rapporti sempre più specifici e frammentati, piegandosi alle esigenze del mercato e della tecnica e rinunciando alla loro funzione di tutela delle differenze e promozione della giustizia. Dal canto suo, il giurista contemporaneo, che opera in questa realtà caotica e priva di riferimenti, è chiamato a scegliere tra l'assecondamento delle regole del mercato e la difesa delle posizioni individuali svantaggiate, che l'ordinamento giuridico, ormai governato dalla tecnica, finisce per marginalizzare.

sperimenta quando è chiamato a misurarsi con l'esigenza di regolamentazione dei sistemi di IA.

Irti propone una visione del diritto contemporaneo caratterizzata da una sovrapproduzione di norme che si susseguono senza comporre alcun disegno organico e rispondono solo a logiche di efficienza e immediata fruibilità. I prodotti della normazione si caratterizzano, così, per la tacita rinuncia all'adesione a valori universali e, pur conformandosi all'osservanza di un iter procedurale prestabilito che ne convalida l'autorità, riflettono esclusivamente il perseguimento di finalità transitorie e parziali.

Nella medesima prospettiva, il progresso tecnologico costante, che spesso si accompagna all'assenza di uno scopo ultimativo e comprensibile – diverso dal mero assecondamento dell'esigenza di sviluppo industriale, strumentale all'incremento del profitto economico – rischia sia di rendere superfluo qualsiasi tentativo di cristallizzazione di tale processo nella forma della regola, sia di far percepire dalla collettività l'intervento del legislatore come un fattore di rallentamento e un ostacolo alla competitività (e, per ciò stesso, alla promozione del benessere).

La norma, intesa come limite obiettivo alle congenite pulsioni prevaricatrici che alimentano l'azione dei centri di potere economico, e il diritto, inteso come parametro di sostenibilità sociale di qualsiasi forma di innovazione, sembrano, così, entrare in conflitto con quegli interessi pubblici (alla sicurezza, alla governabilità, alla celerità delle transazioni) che le nuove tecnologie, e i sistemi di IA in particolare, tendono ad assicurare – in cambio del riconoscimento di uno spazio di operatività privo di condizionamenti e restrizioni.

Il nichilismo giuridico costituisce l'approdo (…inevitabile?[12])

[12] Scrive ancora IRTI, *Nichilismo e metodo giuridico*, in *op. loc. cit.*, 7, "[…] laicizzate le fonti del diritto e sciolto ogni legame con la teologia, le norme sono

del processo di secolarizzazione del diritto positivo e di affrancamento dai suoi fondamenti legittimanti[13]. Con sé reca la consapevolezza che le norme non riflettono più una verità a-storica, di origine divina o naturale, ma rispondono a esigenze contingenti, che coincidono con la volontà espressa dal soggetto detentore del potere e si caratterizzano esclusivamente per la loro conformità al prodotto di una ordinata sequenza procedimentale. La giuridicità si converte in mera tecnica di elaborazione, che esalta l'approccio formalistico a scapito di qualsiasi apprezzamento sostanziale e giudizio di valore. Scaturiscono da ciò l'incremento e la frammentazione delle regole e la stratificazione dei rapporti sottostanti – condizioni che, unite alla mancanza di un orizzonte assiologico condiviso, appaiono in grado di compromettere la capacità del diritto di promuovere la giustizia, intesa come bene universale, e di valorizzare le differenze.

Anche la regolamentazione dei sistemi di intelligenza artificiale risente di tale deriva, testimoniata dalla proliferazione di iniziative

venute nell'esclusivo e totale dominio della volontà umana. I disegni, volti a ricostituire l'unità in forme storiche e mondane, si rivelano ormai falliti. L'età moderna ha esteso al diritto la parola più audace e crudele: 'produrre'. Le norme giuridiche, al pari di qualsiasi bene di mercato, sono 'prodotte': vengono dal nulla e possono esser ricacciate nel nulla. La forza che le 'produce', ossia le chiama innanzi o le rifiuta, le costruisce o le distrugge, è soltanto il volere degli uomini. Le officine giuridiche lavorano in tutte le ore del giorno ed in tutti i luoghi della vecchia Europa: nessuna norma ha privilegio d'immutabilità e d'inviolabilità".

[13] Si può sostenere che il nichilismo giuridico, quale deriva del processo di secolarizzazione del diritto (che conosce nella laicizzazione della tutela dei diritti umani il suo momento culminante: v. T. GUTMANN, *Secolarizzazione del diritto e giustificazione normativa* (trad. it. di P. Chiarello), Napoli, 2016, 13 ss.; E.-W. BÖCKENFÖRDE, *Diritto e secolarizzazione. Dallo Stato moderno all'Europa unita*, a cura di G. PRETEROSSI, Laterza, Roma-Bari, 2007) svela un paradosso: quello della fragilità del precetto normativo, la cui essenza, "liberata" dalla necessità di una legittimazione trascendente, è incapace di fondare solo in sé stessa l'incontestabilità del primato assiologico.

non coerenti tra loro e dalla maggior cura riservata alla classificazione degli ambiti di intervento normativo, anziché alla valutazione delle reali implicazioni etiche e sociali dell'IA e all'approntamento di possibili rimedi[14]. I primi tentativi di normare l'intelligenza artificiale, intesa come nuovo fenomeno sociale, sembrano delinearsi come un insieme di proposte ancora disorganico, inidoneo ad affrontare le questioni fondamentali sollevate dall'impatto di questa tecnologia sull'essere umano e sulla società. D'altronde, lo stesso approccio quantitativo, tipico dell'IA, nel proporre una visione dell'essere umano ridotto ad agglomerato di dati quantificabili e manipolabili, può costituire un fattore potenzialmente in grado di influenzare l'orientamento normativo in materia[15].

Si è accennato al fatto che, a differenza di quanto accade in altri settori, il ritmo di avvicendamento delle disposizioni che dovrebbero regolare i sistemi di IA è dettato dall'andamento del progresso tecnologico – e, quindi, in ultima analisi, dalla stessa capacità dell'intelligenza artificiale di auto-svilupparsi in modi non necessariamente prevedibili. In uno scenario futuro non implausibile, i sistemi di intelligenza artificiale potrebbero, quindi, essere interpellati direttamente, per stabilire quali siano le regole più appropriate per non ostacolarne l'attività. Dietro l'apparente "neutralità" degli algoritmi, tuttavia, si cela la progressiva deresponsabilizzazione del legislatore umano e la possibilità di un suo esautora-

[14] C. D. SCHMIT, M. J. DOERR E J. K. WAGNER, *Leveraging IP for AI governance*, in *Science*, 17 febbraio 2023, 646-648, hanno ipotizzato l'adozione di uno strumento regolatore di natura pubblica basato sul riconoscimento dei titoli di proprietà intellettuale rilasciati per alcuni oggetti e applicazioni dell'IA (ad esempio, modelli di IA e set di dati di addestramento). Secondo questo modello, l'IA potrebbe essere concessa in licenza a condizione che il soggetto titolare aderisca a specifiche pratiche etiche e rispetti un codice di condotta.

[15] Sulla frammentarietà delle iniziative di regolamentazione internazionale dell'IA si tornerà più diffusamente *infra*, par. 3.

mento dall'esercizio delle necessarie forme di controllo di fronte a scelte eticamente discutibili.

Lo sviluppo tecnologico reca con sé il rischio di una deriva nichilista anche nell'ambito delle relazioni giuridiche internazionali. In proposito, si possono distinguere almeno due diversi scenari. Per un verso, l'automazione dei processi decisionali di natura militare ha contribuito ad accrescere il rischio di conflitti armati, per il fatto di rendere più semplice, impersonale e sicuro il ricorso all'uso della forza, esasperandone il potenziale distruttivo. L'IA può essere impiegata, inoltre, per compiere cyberattacchi e, sul versante opposto, per allestire sistemi di sorveglianza e controllo sempre più invasivi[16], che erodono la fiducia reciproca tra gli Stati e compromettono lo sviluppo della cooperazione multilaterale – presupposto indispensabile per la soluzione di problemi globali.

A fronte di ciò, la diffusione di informazioni non verificate e l'incremento della propaganda ostile attraverso l'IA hanno ridotto gli spazi tradizionalmente occupati dal confronto diplomatico, favorendo l'instaurazione di un clima di tensione costante. La disparità di accesso alle tecnologie necessarie allo sviluppo dell'IA determina, inoltre, una progressiva polarizzazione delle posizioni all'interno della comunità internazionale, alla quale si accompagnano una regressione dello spirito di solidarietà tra popoli e il progressivo rifiuto dei valori e dei principi ai quali il diritto internazionale si è ispira.

In ultima analisi, l'avvento dell'IA sembra aver favorito l'instaurazione di un assetto competitivo che, nello spazio di pochi

[16] Si pensi al dibattito giuridico sviluppatosi intorno alla legittimità delle tecnologie di riconoscimento facciale in rapporto alla tutela internazionale dei diritti umani e alle garanzie costituzionali. V. da ultimo, R. DÉCHAUX, *Constitutional Issues of The Use of Facial Recognition Technology*, in M. G. RODOMONTE, L. DURST (*eds.*), *Judicial Review, Fundamental Rights and Rule of Law*, Routledge, London, 2025, 171–182.

anni, ha sostituito il modello cooperativo affermatosi dopo la fine della guerra fredda. È rispetto a tale mutamento che occorre misurare l'efficacia degli strumenti di regolamentazione promossi sul piano internazionale[17].

3. I primi tentativi di regolamentazione internazionale e l'assenza di un approccio globale condiviso

Nel diritto internazionale, i sistemi di IA influenzano sia l'interpretazione e l'applicazione delle norme vigenti, sia i processi di formazione di quelle future. Ad esempio, le regole in materia di commercio internazionale, che accoglievano tradizionalmente la distinzione tra beni e servizi, attraversano una fase di incertezza interpretativa per via della qualificazione dei prodotti del commercio digitale e dell'IA, rispetto ai quali tale distinzione risulta inadeguata.

Più in generale, la capacità dei sistemi di IA di attribuire un crescente grado di autonomia a qualsiasi tipo di infrastruttura informatica mette in guardia sul rischio di sottrarre al controllo

[17] È indicativo il fatto che non vi sia ancora un contributo di riflessione sufficientemente netto sul ruolo che l'IA può svolgere per il potenziamento degli strumenti di tutela dei diritti umani (anche a carattere meramente strumentale, come nel caso dell'impiego delle tecnologie satellitari per il monitoraggio di crisi umanitarie), a testimonianza della difficoltà di considerare tale tecnologia realmente "neutra", in assenza di regole chiare sul relativo impiego. Per alcune considerazioni sul punto si rinvia a S. SHIRAZYAN, M. SISSONS, *AI's Potential to Advance Human Rights? Striking the Right Balance*, in «Just Security», post del 2 agosto 2024 (https://www.justsecurity.org/98097/ais-potential-to-advance-human-rights/, data di accesso 20 marzo 2025). D'altro canto, come evidenziato da J. LEE, *Artificial Intelligence and International Law*, Springer, Cham, 2022, 37 ss., i benefici pratici misurabili in punto di incremento del livello di benessere individuale derivanti dall'avvento dell'IA sono innegabili.

umano il relativo funzionamento e le stesse modalità di comprensione dei processi decisionali[18]. Dinanzi a questa prospettiva, qualsiasi disciplina che intenda limitare le conseguenze pregiudizievoli derivanti da un uso scriteriato dell'IA è chiamata a confrontarsi con le regole e i principi posti a tutela dei diritti umani e del rule of law come parametro fondamentale di legittimità.

È in questa cornice che devono essere valutati i primi tentativi di regolamentazione, anche a carattere non vincolante, che sono stati promossi in ambito internazionale, sia a livello universale, sia regionale. Le Nazioni Unite, attraverso l'Organismo consultivo di alto livello (UN-HLAB) promosso dal Segretario Generale[19] hanno da subito posto l'accento sulla necessità che tutte le fasi legate al ciclo di vita dell'IA – progettazione, sviluppo, diffusione e applicazione – siano concepite e realizzate a beneficio dell'intera umanità[20].

[18] Il cd. "enigma della scatola nera" pone il problema dell'opacità dei processi cognitivi su cui i sistemi di IA basano le proprie scelte. Occorre interrogarsi, infatti, sulle modalità di comprensione di tali processi e su eventuali pregiudizi e disallineamenti suscettibili di determinare esiti imprevedibili. Cfr. T. WISCHMEYER, *Artificial Intelligence and Transparency: Opening the Black Box*, in T. WISCHMEYER, T. RADEMACHER (*eds.*), *Regulating Artificial Intelligence*, Springer, Cham, 2020, 75-101.

[19] Per un approfondimento sui compiti dell'UN-HLAB si rinvia al portale https://www.un.org/digital-emerging-technologies/ai-advisory-body.

[20] L'Assemblea Generale delle NU ha adottato due importanti risoluzioni in materia. La prima, dell'11 marzo 2024, è intitolata "Seizing the Opportunities of Safe, Secure and Trustworthy Artificial Intelligence Systems for Sustainable Development" (UNDOC.A/RES/78/165) e pone l'accento sulla necessità che "[…] human rights and fundamental freedoms must be respected, protected and promoted throughout the life cycle of artificial intelligence systems", esortando gli Stati membri e, "where applicable, other stakeholders" a evitare o abbandonare qualsiasi impiego dei sistemi di IA "that are impossible to operate in compliance with international human rights law or that pose undue risks to the enjoyment of human rights, especially of those who are in vulnerable situations". In modo an-

Nel giugno 2023, il Segretario generale delle Nazioni Unite, António Guterres, ha sostenuto l'idea di un'Agenzia internazionale per l'IA, legata alla nuova "Agenda for Peace". In tale occasione, è stata avviata una riflessione sul modello istituzionale da seguire e sulle funzioni ad essa attribuibili, che vanno dalla valutazione delle capacità attuali e potenziali dei sistemi di IA, per consentire un'effettiva comprensione delle relative implicazioni, all'assistenza agli Stati membri nell'implementazione di standard comuni nei rispettivi ordinamenti – con particolare attenzione alla legislazione dei paesi in via di sviluppo, per prepararli ad affrontare gli impatti differenziati dell'IA nei loro sistemi nazionali.

L'idea è stata ripresa dall'UN-HLAB nel rapporto finale "Governing AI for Humanity", presentato il 19 settembre 2024 all'Assemblea Generale delle Nazioni Unite[21]. Il rapporto si sofferma sui rischi legati allo sviluppo non controllato dell'IA e al suo potenziale trasformativo a livello globale, sottolineando l'assenza di un coordinamento internazionale e la frammentarietà degli strumenti di regolamentazione finora promossi. Le sette raccomandazioni finali sono volte a colmare le lacune esistenti nella go-

cor più significativo, la risoluzione "reaffirms that the same rights that people have offline must also be protected online, including throughout the life cycle of artificial intelligence systems" (par. 5). La seconda, intitolata "Enhancing International Cooperation on Capacity-Building of Artificial Intelligence", è stata adottata il successivo 25 giugno (UNDOC.A/RES/78/311).

[21] Il rapporto è reperibile qui: https://www.un.org/sites/un2.un.org/files/governing_ai_for_humanity_final_report_en.pdf.

Una delle iniziative più significative intraprese dall'UN-HLAB è stata la promozione di un ampio numero di consultazioni globali, per dar vita a un processo inclusivo e partecipativo, finalizzato a raccogliere opinioni, prospettive e competenze da una vasta gamma di soggetti interessati. Le consultazioni hanno consentito di collegare tra loro le diverse prospettive e promuovere un dialogo costruttivo tra i vari gruppi di stakeholder coinvolti nello sviluppo e nell'implementazione dei sistemi di IA.

vernance globale dell'IA e a promuovere una cooperazione internazionale più efficace, nel rispetto di principi fondamentali come l'inclusività, l'interesse pubblico, la gestione dei dati, l'universalità, la collaborazione e l'osservanza del diritto internazionale.

Meritano, in proposito, di essere richiamate la proposta di istituzione di un Panel scientifico internazionale, a carattere indipendente e multidisciplinare, che, attraverso l'adozione di rapporti annuali sulle capacità, le opportunità, i rischi e le incertezze dell'IA, e la pubblicazione di analisi tematiche e rapporti ad hoc su questioni emergenti, persegua il risultato di una comprensione scientifica generalizzata dell'IA e delle sue applicazioni, colmando il divario informativo sussistente tra i paesi leader nello sviluppo di tale tecnologia e il resto del mondo[22].

All'istituzione di questo organismo scientifico indipendente dovrebbero accompagnarsi l'avvio di un dialogo politico semestrale sulla governance dell'IA presso le Nazioni Unite, finalizzato alla condivisione di esperienze e buone prassi e – su base volontaria – di incidenti significativi legati all'impiego dell'IA e l'istituzione di una piattaforma che riunisca rappresentanti di organizzazioni per lo sviluppo di standard, aziende tecnologiche, società civile e il panel scientifico, con lo scopo di realizzare una tassonomia dedicata all'IA, periodicamente aggiornata[23].

Nel rapporto si propone anche l'istituzione di un fondo globale, gestito in modo indipendente, per la gestione di contributi finanziari e materiali preordinati a semplificare l'accesso ai fattori

[22] Sottolinea il rapporto come "[t]here is a need for timely, impartial and reliable scientifc knowledge and information about AI for Member States to build a shared foundational understanding worldwide, and to balance information asymmetries between companies housing expensive AI labs and the rest of the world, including via information-sharing between AI companies and the broader AI community" (*ibid.*, 48, par. 89).

[23] *Ibid*, 52 ss.

abilitanti dell'IA nei paesi in via di sviluppo. Al contempo, si raccomanda la definizione di un quadro regolatorio generale, promosso da un organismo come la Commissione delle Nazioni Unite per il Diritto Commerciale Internazionale (UNCITRAL), che stabilisca principi e regole comuni per la governance dei dati di addestramento dell'IA e istituisca meccanismi uniformi di gestione e scambio dei dati per modellare il mercato[24].

Al di là dell'organismo consultivo promosso in ambito ONU, occorre dar conto di un variegato ecosistema di iniziative multilaterali o di portata regionale, tutte orientate a promuovere standard di applicazione volontaria sulla governance dell'IA, *lato sensu* intesa.

Al riguardo, vanno menzionati anzitutto i Principi sull'IA dell'Organizzazione per la Cooperazione e lo Sviluppo Economico (OCSE), originariamente adottati nel 2019 e aggiornati nel 2024. Si tratta del primo standard intergovernativo di regolamentazione dell'IA, volto a promuovere sistemi e applicazioni affidabili, che rispettino i diritti umani e i valori democratici[25]. Sembra utile richiamare, altresì, il "Processo di Hiroshima" sull'IA promosso dal G7[26]; le Linee Guida del G20[27]; le Linee guida per l'IA

[24] *Ibid*, p. 65 ss.

[25] Lo standard dell'OCSE sull'IA si compone di cinque principi basati su valori e cinque raccomandazioni che forniscono una guida pratica e flessibile per i responsabili politici e gli attori coinvolti nelle attività di sviluppo, gestione e implementazione dei sistemi di IA. Il documento è reperibile qui:
https://www.oecd.org/en/topics/sub-issues/ai-
princi-
ples.html#:~:text=The%20OECD%20AI%20Principles%20promote,stand%20the%20test%20of%20time.

[26] Il cd. "Processo di Hiroshima" sull'IA è stato promosso nel maggio 2023 sulla base delle indicazioni dei leader riuniti al vertice G7 di Hiroshima sotto la presidenza del Giappone. Nel mese di dicembre, gli Stati partecipanti hanno raggiunto un accordo su un primo quadro comune di iniziative, noto come "Hiro-

dell'Associazione delle Nazioni del Sud-Est Asiatico[28]; la Strategia continentale per l'IA dell'Unione Africana[29]; la Dichiarazione di Nuova Delhi del Partenariato Globale sull'IA[30]; la Dichiarazione di Bletchley sulla sicurezza dell'IA[31]; la Dichiarazione di Seoul per

shima AI Process Comprehensive Policy Framework". Il quadro si compone di una serie di misure, tra cui i "Principi guida internazionali del processo di Hiroshima per tutti gli attori dell'IA" e il "Codice di condotta internazionale del processo di Hiroshima per le organizzazioni che sviluppano sistemi avanzati di IA". I Principi guida costituiscono l'insieme dei principi che dovrebbero essere applicati a tutti gli attori coinvolti nel ciclo di vita dell'IA, come nel caso degli sviluppatori, chiamati a segnalare pubblicamente le capacità in possesso dei sistemi avanzati di IA e gli ambiti di utilizzo inappropriato. Anche il Codice di condotta si rivolge ai medesimi soggetti, elencando in modo più dettagliato talune azioni da intraprendere per garantire lo sviluppo e la diffusione di tecnologie che consentano agli utenti di identificare con precisione i contenuti generati dall'intelligenza artificiale (https://www.mofa.go.jp/files/100573473.pdf).

[27] Per un approfondimento si rinvia a https://www.caidp.org/resources/g20/#:~:text=The%20G20%20guidelines%20call%20for,and%20internationally%20recognized%20labor%20rights.

[28] Il documento è reperibile qui: https://asean.org/wp-content/uploads/2024/02/ASEAN-Guide-on-AI-Governance-and-Ethics_beautified_201223_v2.pdf.

[29] È significativo che l'approccio promosso dall'Unione Africana in materia associ l'impiego dei sistemi di IA alla crescita culturale e allo sviluppo sostenibile degli Stati del continente. Nel documento si legge: "[t]he Mission of the AI Strategy is to harness AI for accelerating social and economic transformation and promoting cultural renaissance in Africa in line with the AU Agenda 2063 and the Sustainable Development Goals, to minimise the risks that AI poses to African people and countries, and to accelerate the development of the necessary AI capabilities of the AU Member States" (si veda https://au.int/sites/default/files/documents/44004-doc-EN-_Continental_AI_Strategy_July_2024.pdf).

[30] Il documento è reperibile qui: https://gpai.ai/gpai-new-delhi-declaration-2024.pdf.

[31] La Dichiarazione di Bletchley è stata adottata all'esito del Summit sulla sicurezza dell'IA tenutosi nel novembre 2023. Per un approfondimento sui relativi contenuti v.: https://www.gov.uk/government/publications/ai-safety-summit-

un'IA sicura, innovativa e inclusiva[32]; nonché diversi strumenti etici, tra cui la Dichiarazione dell'UNESCO sull'etica dell'IA[33]; la Dichiarazione di Santiago per promuovere l'IA etica in America Latina e nei Caraibi[34] e l'Appello di Roma per l'etica dell'IA[35].

Benché tale mosaico di proposte, nella sua ampiezza ed eterogeneità, possa apprezzarsi come un elemento in sé positivo, che riflette la comune esigenza di governare il progresso tecnologico e le modalità di applicazione dell'IA nei suoi diversi ambiti di operatività, l'evidente frammentarietà e l'assenza di un coordinamento unitario pongono un problema di efficacia. È appena il caso di osservare che il ricorso a misure parziali e non inclusive, e pertanto non esenti dal rischio di sovrapposizioni, contraddizioni o lacune rispetto a quanto previsto in strumenti omologhi, reca con sé un evidente vulnus applicativo, suscettibile di essere sfruttato per sco-

2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023.

[32] Alla Dichiarazione è associato uno "Statement of Intent toward International Cooperation on AI Safety Science", in cui si sottolinea "the role of scientific inquiry and the benefits of international coordination for the advancement of such inquiry, so that ultimately the benefits of AI development and deployment are shared equitably around the globe" (par. 2.1). Entrambi i documenti sono reperibili qui: https://www.gov.uk/government/publications/seoul-declaration-for-safe-innovative-and-inclusive-ai-ai-seoul-summit-2024.

[33] Il testo è reperibile qui: https://unesdoc.unesco.org/ark:/48223/pf0000381137.

[34] Il testo è reperibile qui: https://minciencia.gob.cl/uploads/filer_public/40/2a/402a35a0-1222-4dab-b090-5c81bbf34237/declaracion_de_santiago.pdf.

[35] L'Appello di Roma propone l'espressione "algorethics", per significare "the development of an artificial intelligence that serves every person and humanity as a whole; that respects the dignity of the human person, so that every individual can benefit from the advances of technology; and that does not have as its sole goal greater profit or the gradual replacement of people in the workplace" (https://www.romecall.org/).

pi malevoli e per compromettere gli sforzi prodotti dagli attori (pubblici e privati) più diligenti.

Appare, altresì, superfluo soffermarsi sugli effetti intrinsecamente transfrontalieri dei sistemi di IA, la cui regolamentazione esige una risposta globale coerente e armonizzata, adattabile a diversi contesti e rappresentativa delle diverse prospettive emergenti a livello regionale. Tuttavia, non si registra ancora un sufficiente livello di consapevolezza circa l'importanza di una partecipazione attiva e diffusa all'adozione di strumenti condivisi, come dimostrato dalle critiche all'eccesso di regolamentazione attribuito all'Unione europea dopo l'adozione dell'AI Act.

Due fattori in particolare sembrano alimentare la segnalata frammentazione, impedendo l'avanzamento del processo di elaborazione di una normativa internazionale a vocazione universale. Da un lato, il carattere dual use di numerose applicazioni dell'IA rende impossibile, o comunque altamente complesso, dissociare formalmente i sistemi impiegati in ambito civile da quelli utilizzati in campo militare. Dall'altro, lo sviluppo e l'implementazione di tali sistemi costituiscono attività generalmente ricondotte nel perimetro della sicurezza nazionale, e quindi per definizione insuscettibili di essere assoggettate a un regime giuridico condiviso con altri Stati. Anche il primo strumento internazionale a carattere vincolante adottato in materia subisce alcuni condizionamenti sul piano applicativo, proprio in ragione del rilievo attribuito a tali fattori, con ovvie ricadute in punto di efficacia.

4. La Convenzione quadro del Consiglio d'Europa sull'intelligenza artificiale, i diritti umani, la democrazia e lo Stato di diritto tra ambizioni e lacune

Il 5 settembre 2024 è stata conclusa a Vilnius la Convenzione

quadro del Consiglio d'Europa sull'intelligenza artificiale, i diritti umani, la democrazia e lo Stato di diritto[36]. Oltre ai rappresentanti dei 46 Stati membri del CoE, hanno partecipato alla fase negoziale anche altri Stati (Australia, Argentina, Canada, Costa Rica, Israele, Giappone, Messico, Perù, Stati Uniti d'America e Uruguay) ed entità non statali (UE, Santa Sede), legittimate, proprio a motivo della loro partecipazione, ad acquisire lo status di parte[37]. Altri Stati non membri del CoE potranno essere invitati ad aderire alla Convenzione dopo la sua entrata in vigore, previo consenso unanime delle parti e il parere favorevole del Comitato dei Ministri del Consiglio d'Europa.

In conformità all'articolo 1, la Convenzione disegna una regime di regole volto a garantire che tutte le attività "within the life-

[36] Già il Principio 1.2 elaborato dall'OCSE, e intitolato "Respect for the rule of law, human rights and democratic values, including fairness and privacy" si concentra sulla necessità che "AI actors should respect the rule of law, human rights, democratic and human-centred values throughout the AI system lifecycle", tra cui "non-discrimination and equality, freedom, dignity, autonomy of individuals, privacy and data protection, diversity, fairness, social justice, and internationally recognised labour rights". Il rispetto di tali principi e valori implica anche "addressing misinformation and disinformation amplified by AI, while respecting freedom of expression and other rights and freedoms protected by applicable international law". A tal fine, "AI actors should implement mechanisms and safeguards, such as capacity for human agency and oversight, including to address risks arising from uses outside of intended purpose, intentional misuse, or unintentional misuse in a manner appropriate to the context and consistent with the state of the art". Come si vedrà, il tema del controllo è ripreso anche nella Convenzione del Consiglio d'Europa.

[37] Il testo finale della Convenzione è stato adottato il 17 maggio 2024. Tra i primi firmatari figurano l'Unione europea (UE), gli Stati Uniti, il Regno Unito, Israele, Norvegia, Georgia, Moldavia, Islanda, Andorra e San Marino. L'entrata in vigore della Convenzione è subordinata al deposito di cinque strumenti di ratifica, di cui almeno tre da parte di Stati membri del Consiglio d'Europa. Lo stato delle ratifiche è consultabile qui: https://www.coe.int/en/web/conventions/full-list?module=signatures-by-treaty&treatynum=225.

cycle of artificial intelligence systems"[38] nella disponibilità degli Stati parti rispettino i diritti umani, la democrazia e lo Stato di diritto. Si tratta della disposizione centrale dell'accordo, la cui formulazione pone giusta enfasi sulla necessità che tutte le fasi e le modalità applicative dei sistemi di IA siano coerenti con i valori del costituzionalismo democratico[39]. Sembra superfluo osservare che, nella congiuntura attuale, tale compatibilità appare tutt'altro che scontata e riflette un approccio che potrebbe dissuadere alcuni paesi firmatari dall'intraprendere l'iter di ratifica.

Per conseguire le finalità di cui all'articolo 1, gli Stati parti sono chiamati ad adottare o a mantenere in vigore norme legislative, amministrative o di altro genere, idonee a dare attuazione alle disposizioni della Convenzione, incluse eventuali misure di natura specifica, ovvero a carattere trasversale, applicabili cioè indipendentemente dal tipo di tecnologia utilizzata. La scelta in ordine alla misura da adottare dipenderà dalla gravità della situazione (vale a dire che la legislazione nazionale dovrebbe prevedere una gradazione di intensità delle norme applicabili in determinate circostan-

[38] L'articolo 2 della Convenzione definisce "artificial intelligence system" come "a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations or decisions that may influence physical or virtual environments. Different artificial intelligence systems vary in their levels of autonomy and adaptiveness after deployment".

[39] Tra i motivi di preoccupazione, evidenziati nel Preambolo, che hanno reso opportuna l'adozione di una disciplina internazionale a carattere vincolante, figurano anche "the misuse of artificial intelligence systems" e "the use of such systems for repressive purposes in violation of international human rights law, including through arbitrary or unlawful surveillance and censorship practices that erode privacy and individual autonomy". Si vedrà, tuttavia, che le ampie deroghe previste nel testo della Convenzione per il perseguimento di interessi legati alla sicurezza nazionale appaiono in grado di compromettere in radice l'efficacia dell'azione di controllo demandata all'accordo.

ze) e dalla maggiore o minore probabilità che, durante il ciclo di vita di un sistema di IA, possano verificarsi effetti pregiudizievoli sul rispetto dei diritti umani o sulla tenuta dello Stato di diritto[40]. Non è chiaro, peraltro, quali interventi possano farsi rientrare nella categoria residuale delle "other measures", tenuto conto che le norme legislative e quelle amministrative dovrebbero esaurire da sole l'intera gamma delle misure vincolanti – a meno di non ritenere che i redattori dell'accordo abbiano inteso così valorizzare il ricorso a strumenti di soft law.

L'ambito di applicazione della Convenzione, stabilito nel successivo articolo 3, copre tutte le attività del ciclo di vita dei sistemi di IA suscettibili di interferire con i diritti umani, la democrazia e lo Stato di diritto, sia mediante l'adozione di misure che intervengano sulle attività (a) intraprese da autorità pubbliche o da soggetti privati che agiscono per loro conto, sia (b) affrontando i rischi e le conseguenze derivanti dalle medesime attività, qualora intraprese da attori privati ma escluse dalla portata applicativa dell'ipotesi precedente, in modo conforme all'oggetto e allo scopo dell'accordo.

A tal fine, gli Stati parti sono tenuti a specificare, all'interno di una dichiarazione presentata al Segretario Generale del Consiglio d'Europa al momento della firma o del deposito dello strumento di ratifica, le modalità con cui intendono dare attuazione ai principi e agli obblighi enunciati nei Capitoli da II a VI della Convenzione in rapporto alle attività promosse da attori privati, ovvero a

[40] Il testo del par. 2 dell'articolo 1 recita: "[e]ach Party shall adopt or maintain appropriate legislative, administrative or other measures to give effect to the provisions set out in this Convention. These measures shall be graduated and differentiated as may be necessary in view of the severity and probability of the occurrence of adverse impacts on human rights, democracy and the rule of law throughout the lifecycle of artificial intelligence systems. This may include specific or horizontal measures that apply irrespective of the type of technology used".

proporre l'adozione di altre misure, ritenute maggiormente appropriate[41].

Suscita alcune perplessità la formulazione del paragrafo 2, che esclude dall'ambito di applicazione dell'accordo tutte le attività riconducibili al ciclo di vita dei sistemi di IA laddove inerenti alla tutela "of its national security interests", purché lo svolgimento di tali attività risulti compatibile "with applicable international law, including international human rights law obligations, and with respect for its democratic institutions and processes".

Si tratta di una condizione molto specifica e di importanza decisiva nell'economia della convenzione, tenuto conto della particolare delicatezza del settore in cui i sistemi di IA sono chiamati a operare. Escludere dalla portata applicativa della convenzione la tutela degli interessi legati alla sicurezza nazionale rischia, infatti, di compromettere il raggiungimento degli stessi scopi per cui l'accordo è stato concepito, data l'ampiezza di una simile deroga e l'elevata discrezionalità di cui gli Stati parti godrebbero nel ricondurre nell'alveo della sicurezza nazionale lo sviluppo e la gestione di talune applicazioni.

Sul punto, può osservarsi che la tutela della sicurezza nazionale rappresenta di per sé un ambito potenzialmente confliggente con il rispetto degli obblighi internazionali in materia di diritti umani e relativi "democratic institutions and processes", giacché il perseguimento di interessi pubblici nel settore della sicurezza nazionale

[41] La disposizione prosegue precisando che nell'esecuzione di tali adempimenti, le parti non possono derogare o limitare l'applicazione dei propri obblighi internazionali assunti per proteggere i diritti umani, la democrazia e lo Stato di diritto. Al momento, solo la Norvegia ha presentato tale Dichiarazione, affermando che "it shall apply the principles and obligations set forth in Chapters II to VI of this Convention to activities of private actors" (il testo della Dichiarazione è disponibile qui: https://www.coe.int/en/web/conventions/full-list?module=declarations-by-treaty&numSte=225&codeNature=0).

non di rado esige una limitazione delle ordinarie modalità di esercizio dei diritti fondamentali e di accesso alle istituzioni deputate alla loro protezione. Basti solo pensare alle modalità di impiego dei sistemi di IA in materia di sorveglianza e controllo per il mantenimento dell'ordine pubblico e al rischio di violazioni nella sfera di autonomia privata che il ricorso a tali tecnologie è in grado di paventare.

La previsione di un'esenzione generalizzata per il settore della sicurezza nazionale introduce, pertanto, un'incrinatura non trascurabile, suscettibile di minare la stessa ratio ispiratrice del trattato. Se può ritenersi comprensibile la volontà degli Stati di preservare la propria autonomia in un settore fondamentale per il nuovo assetto competitivo che contraddistingue la comunità internazionale in questa fase storica, una deroga di così ampia portata rischia di creare una zona franca normativa, sottraendo all'esigenza di tutela dei diritti fondamentali proprio quelle applicazioni dell'IA che, per la loro natura intrusiva e per il potenziale impatto sulle libertà individuali e collettive, esigerebbero la subordinazione a forme di controllo più rigorose di quelle ordinarie.

Ciò che suscita maggiore perplessità è l'opacità che inevitabilmente connota i sistemi di IA operanti nel settore della sicurezza nazionale. Tale caratteristica induce a ritenere che, una volta esentati dall'osservanza degli obblighi di trasparenza e conformità al rispetto dei diritti umani e del rule of law, tali sistemi potrebbero essere impiegati per scopi di sorveglianza massiva e profilazione individuale, favorendo derive autoritarie e la progressiva compressione delle libertà fondamentali. Parimenti, l'esenzione generalizzata di un settore così cruciale suggerisce un messaggio fuorviante in punto di sviluppo di una cultura globale della responsabilità nel settore dell'IA, legittimando un approccio basato su logiche di potenza e di contrapposizione.

Come suggerito da diversi stakeholder[42], per la stessa credibilità del regime di regole promosso dalla Convenzione, è auspicabile che gli Stati parti possano riconsiderare la portata di tale esenzione, privilegiando un approccio più mirato, in linea con l'impostazione delle norme della CEDU, che – com'è noto – riconosce la possibilità di limitare il godimento di alcuni diritti (in particolare quelli di cui agli articoli da 8 a 11) per ragioni di sicurezza nazionale, a condizioni che le restrizioni imposte dalle autorità non risultino sproporzionate in rapporto allo scopo perseguito[43].

Dal tenore della previsione di cui si discute, non è chiaro, inoltre, a chi spetti la valutazione circa la compatibilità tra i sistemi di IA utilizzati per finalità di sicurezza nazionale e i principi democratici; se cioè un simile accertamento rimanga – come parrebbe – di esclusivo appannaggio dello Stato o sia, invece, sindacabile, quanto meno sul piano politico, dalla conferenza delle parti o da altro soggetto terzo (ad esempio, il Comitato dei Ministri del Consiglio d'Europa) legittimato a valutare il rispetto della Convenzione da parte degli Stati contraenti. Né, infine, è dato di comprendere quali siano le conseguenze derivanti da un eventuale accertamento ne-

[42] Si veda la lettera aperta inviata il 5 marzo 2024 da un gruppo di cittadini, membri di organizzazioni della società civile, istituzioni accademiche e di ricerca, esperti di intelligenza artificiale e tecnologie digitali agli Stati membri del Consiglio d'Europa che hanno negoziato la Convenzione. Il documento, intitolato "Do not water down our rights", esortava i negoziatori a includere nel testo della Convenzione le previsioni necessarie "(i) to cover equally the public and private sectors" e "(ii) to unequivocally reject blanket exemptions regarding national security and defence" (https://algorithmwatch.org/de/wp-content/uploads/2024/03/Open_letter_ Council_of_Europe_AI_Convention. pdf).

[43] Per una ricognizione della giurisprudenza della Corte europea dei diritti dell'uomo in materia di tutela dei diritti umani e protezione della sicurezza nazionale si rinvia alla rassegna "National security and European case-law" elaborata dal Servizio ricerche della Corte e risalente al 2013: https://rm.coe.int/168067d214.

gativo. In linea di principio, vista la collocazione della fattispecie tra le disposizioni di cui all'articolo 3, interamente dedicato alla delimitazione dell'ambito di applicazione dell'accordo, si sarebbe portati a sostenere che, nel caso in cui l'impiego di un sistema di IA per finalità di sicurezza nazionale dovesse denunciare profili di incompatibilità con gli obblighi di rispetto dei diritti umani e dei principi democratici, il regime di deroga dovrebbe ritenersi inoperante, con la conseguenza di estendere il regime convenzionale di controllo anche ai suddetti sistemi.

Tuttavia, un esito del genere non trova esplicita conferma nel testo della disposizione, né sembra avvalorato dal contenuto dell'obbligo gravante sugli Stati parti in forza dell'articolo 24 della Convenzione. Questa previsione, infatti, assegna alle parti l'onere di trasmettere alla Conferenza delle parti, entro due anni dalla ratifica della Convenzione e, successivamente, con cadenza periodica, un rapporto recante le misure di dettaglio adottate solo "to give effect to Article 3, paragraph 1, sub-paragraphs a and b", senza stabilire alcun adempimento aggiuntivo qualora le condizioni stabilite dal par. 2 non dovessero risultare soddisfatte.

Più utile, a tal fine, potrebbe rivelarsi il contributo del meccanismo previsto dal successivo articolo 26. Tale disposizione impone alle parti di istituire un meccanismo effettivo, indipendente e imparziale per esercitare una funzione di vigilanza sul rispetto degli obblighi convenzionali, ovvero di incaricare uno o più organi interni già esistenti del medesimo compito. L'organo così istituito o incaricato dovrà essere dotato dei poteri, delle competenze e delle risorse necessari per svolgere l'anzidetta funzione di controllo, che la Convenzione associa di preferenza ai compiti di norma già assegnati agli organismi nazionali in materia di tutela e promozione dei diritti umani, con cui il nuovo meccanismo dovrà coordinarsi[44].

[44] Il par. 4 dell'articolo 26 stabilisce un dovere di favorire il coordinamento

In quest'ottica, è ipotizzabile che, qualora dallo sviluppo e/o dall'impiego dei sistemi di IA nel settore della sicurezza nazionale emergano eventuali profili di incompatibilità con gli standard internazionali posti a protezione dei diritti umani e dello Stato di diritto, sia proprio l'organismo indipendente istituito a mente dell'articolo 26 della Convenzione a doversi far carico del dovere di accertamento. Tuttavia, trattandosi di un organismo interno – sia pure creato in forza di una specifica previsione convenzionale – sembra arduo immaginare che possa essere dotato di poteri ispettivi e di vigilanza particolarmente incisivi, specie in un settore così complesso come quello della sicurezza nazionale.

Struttura analoga presenta la disposizione di cui al par. 3 dell'articolo 3, che esclude dall'ambito di applicazione della Convenzione di Vilnius tutte le attività di ricerca e sviluppo riguardanti le applicazioni dell'IA non ancora disponibili per l'uso, a meno che nelle rispettive fasi di sperimentazione sia emersa una potenziale interferenza con la tutela dei diritti umani e il rispetto dei principi di democrazia e stato di diritto. In proposito, è agevole osservare che la previsione in commento, così formulata, preclude l'esercizio di qualsiasi controllo preventivo sulla trasparenza dei processi di sviluppo dei nuovi sistemi di IA in rapporto alle finalità promosse dalla Convenzione, con il risultato che eventuali applicazioni "strutturalmente" non conformi al rispetto dei diritti umani e dei principi democratici potranno essere sperimentate e prodotte, senza poter essere valutate fino al momento del loro effettivo utilizzo. Nella versione originaria del testo, in realtà, questa ipotesi di dero-

tra i meccanismi di controllo operanti in sede nazionale: "[i]f a Party has provided for mechanisms different from existing human rights structures, it shall take measures, where practicable, to promote effective cooperation between the mechanisms referred to in paragraph 1 and those existing domestic human rights structures".

ga non era contemplata e la Convenzione risultava applicabile ai sistemi di IA anche nelle fasi di progettazione e sviluppo. La formulazione attuale rappresenta, pertanto, un tentativo di compromesso, rispetto al quale, però, l'identificazione delle attività potenzialmente dannose – che la disposizione riconduce essenzialmente al "testing or similar activities" – assume contorni incerti e problematici sul piano interpretativo. Valgono, altresì, le considerazioni fatte in precedenza, posto che anche questa disposizione non chiarisce a quale soggetto debba attribuirsi il potere-dovere di indagare sulle interferenze anzidette, né le conseguenze che potrebbero discendere in concreto dal relativo accertamento, in punto di applicabilità della disciplina prevista dalla convenzione.

Senz'altro più chiara sotto il profilo letterale – seppur parimenti gravida di conseguenze potenzialmente pregiudizievoli – appare, invece, la clausola di cui al paragrafo 4 dell'articolo 3, che esclude incondizionatamente le "[m]atters relating to national defence" dalla portata applicativa dell'accordo. Anche l'esenzione generale per i sistemi di IA utilizzati nel settore della difesa nazionale configura una grave lacuna applicativa del regime giuridico stabilito dalla Convenzione. Gli Stati parti, infatti, potrebbero agevolmente avvalersi di tale ulteriore deroga per giustificare condotte lesive dei diritti umani riconducibili all'applicazione di sistemi di IA, il che appare particolarmente allarmante, considerato il già frequente ricorso a tali sistemi per lo sviluppo e la produzione di armi autonome.

La previsione sembra ispirarsi al contenuto dell'articolo 1 (d) dello Statuto del Consiglio d'Europa, a norma del quale "[m]atters relating to national defence do not fall within the scope of the Council of Europe". Tuttavia, mentre la previsione in parola aveva lo scopo di escludere dall'ambito di intervento dell'organizzazione questioni di natura sovrana come le spese per armamenti, le allean-

ze, la condotta delle operazioni militari o gli aspetti politici della difesa nazionale, lo stesso non può dirsi per i sistemi di IA applicabili in tale ambito, che soggiacciono a logiche diverse, in principio non estranee a una valutazione di conformità con i principi democratici. È appena il caso di ricordare, infatti, che i valori ai quali la Convenzione quadro si ispira – e che intende tutelare di fronte alla potenziale minaccia di una tecnologia di IA orientata a legittimare derive autoritarie o a perseguire scopi illiberali – non vengono meno nella loro sostanza di fondamenta del patto sociale qualora un paese sia chiamato a compiere scelte rilevanti per la propria difesa, pur potendo subire opportune limitazioni sul piano applicativo in situazioni di emergenza[45]. Nella circostanza, tuttavia, i redattori della Convenzione hanno seguito una logica diversa, analoga a quella che ha condotto alla previsione di una deroga in materia di sicurezza nazionale, ritenendo cioè che l'assoggettamento dei sistemi di IA sviluppati per il settore della difesa a uno scrutinio di conformità con gli obblighi internazionali in materia di diritti umani e con i principi democratici potesse costituire un condizionamento eccessivo in rapporto all'esigenza di gestire le politiche di difesa nazionale senza vincoli o controlli[46]. D'altro canto, si è già avuto modo di osservare incidentalmente come diverse applicazioni di IA siano caratterizzate da un impiego duale – consi-

[45] Il modello di riferimento è quello stabilito dall'articolo 15, par. 1, della Convenzione europea dei diritti dell'uomo ("Derogation in time of emergency"), a norma del quale "[i]n time of war or other public emergency threatening the life of the nation any High Contracting Party may take measures derogating from its obligations under this Convention to the extent

strictly required by the exigencies of the situation, provided that such measures are not inconsistent with its other obligations under international law".

[46] L'esenzione del settore della difesa dall'ambito di applicazione della convenzione, strenuamente sostenuta dalla Commissione europea durante i negoziati, trova corrispondenza nell'analoga esclusione prevista dal regolamento dell'UE sull'intelligenza artificiale.

derato che buona parte del progresso tecnologico ha avuto origine in contesti militari per poi trovare applicazione anche in campo civile (basti pensare al web). Sottrarre tali applicazioni, per il solo fatto di essere concepite e/o applicate al settore militare, a un test di compatibilità con i diritti umani costituisce, quindi, una grave limitazione della portata materiale della Convenzione.

Definito nell'articolo 3 il campo di applicazione, la Convenzione chiarisce il contenuto degli obblighi generali posti a carico delle parti nei successivi articoli 4 e 5. Il primo prescrive l'adozione (o il divieto di abrogazione, qualora già in vigore) di misure volte a garantire che i sistemi di IA operativi a livello nazionale si conformino agli obblighi in materia di diritti umani assunti dagli Stati parti nel rispetto dei pertinenti strumenti di natura internazionale e delle corrispondenti norme di diritto interno. Il rinvio materiale operato da tale disposizione produce come primo effetto l'assoggettamento dei sistemi di IA all'attività di controllo svolta dalle istituzioni nazionali e internazionali in materia di diritti umani, affinché l'impiego dei suddetti sistemi non sia motivato da finalità elusive dei pertinenti obblighi di tutela e promozione gravanti sulle autorità pubbliche. Mediante l'adozione delle misure imposte dall'articolo 4 si realizza, pertanto, una piena assunzione di responsabilità dello Stato circa il rispetto dei diritti umani da parte delle applicazioni che caratterizzano il ciclo di vita dei sistemi di IA, anche quando dette applicazioni siano sviluppate o gestite da soggetti privati.

Maggiore rilievo, per il perseguimento delle finalità generali individuate dalla Convenzione, presentano gli obblighi di cui al successivo articolo 5, che impongono agli Stati parti l'adozione di ulteriori misure volte, da un lato, ad assicurare che i sistemi di IA non siano utilizzati per minare l'integrità, l'indipendenza e l'efficacia delle istituzioni e dei processi democratici, "including

the principle of the separation of powers, respect for judicial independence and access to justice"[47] e, dall'altro, a proteggere i processi democratici interni nel contesto delle attività riconducibili al ciclo di vita dei sistemi di IA, incluso il libero accesso e la partecipazione degli individui al dibattito pubblico e la capacità di formarsi opinioni libere.

Gli articoli da 6 a 13 della Convenzione richiamano alcuni principi fondamentali che gli Stati parti sono chiamati ad applicare in fase di implementazione dei sistemi di IA. Si tratta, segnatamente, dei principi di dignità umana e autonomia individuale (art. 7), trasparenza e controllo (art. 8), responsabilità (art. 9, nella duplice veste dell'accountability e della responsibility), uguaglianza e non discriminazione (art. 10), tutela della privacy e dei dati personali (art. 11), affidabilità (art. 12) e innovazione sicura (art. 13).

Ai fini della corretta implementazione del regime di tutela stabilito dalla Convenzione assumono specifico rilievo le disposizioni di cui agli articoli 14 e 15, dedicate ai rimedi e alle garanzie procedurali. In particolare, il primo dei due articoli impone agli Stati parti di adottare o mantenere in vigore tutte le misure necessarie a garantire l'accesso a rimedi efficaci per le violazioni dei diritti umani provocate dalle attività riconducibili al ciclo di vita dei sistemi di IA (par. 1), incluse quelle aventi finalità strumentale (quali la garanzia di accesso, da parte degli organismi autorizzati e dei soggetti interessati, alle informazioni rilevanti relative ai sistemi di IA potenzialmente interferenti con il godimento dei diritti umani; l'ulteriore garanzia circa il fatto che le predette informazioni siano sufficienti per contestare eventuali decisioni o le modalità di impiego del sistema stesso; infine, la garanzia per i soggetti interessati di presentare un reclamo alle autorità competenti). Sugli Stati parti grava anche l'obbligo di assicurare le opportune garanzie procedu-

[47] Così il par. 1 dell'articolo 5 della Convenzione.

rali, tutele e diritti, in conformità con il diritto internazionale e interno applicabile, per le persone le cui prerogative siano state lese dalle attività di un sistema di IA (articolo 15, par. 1).

Completano l'impianto normativo della Convenzione i Capitoli V (articolo 16), dedicato alla gestione dei rischi, VI (articoli 17-22), relativo all'attuazione dell'accordo (che include disposizioni come l'articolo 18, che riguarda i diritti dei bambini e delle persone con disabilità, o l'articolo 20 sull'alfabetizzazione digitale e le competenze digitali) e VII (articoli 23-26), recante i meccanismi di follow-up e di cooperazione, con la già richiamata previsione di un meccanismo di controllo obbligatorio (articolo 26).

5. Considerazioni conclusive: i modelli di regolamentazione europei come argine al primato della tecnica.

L'adozione della Convenzione di Vilnius rappresenta certamente un passo importante per garantire che le tecnologie dell'IA vengano sviluppate e utilizzate in modo conforme al rispetto dei diritti umani e delle regole democratiche. La cornice dei principi ivi richiamati costituisce una difesa fondamentale per la società civile innanzi al rischio di possibili abusi e fornisce un parametro di legittimità con cui i sistemi di IA saranno chiamati a confrontarsi, quando – auspicabilmente in tempi brevi – l'accordo entrerà in vigore. In linea di principio, la sua applicazione rigorosa, da parte di un numero significativo di Stati, dovrebbe condurre a una selezione dei sistemi di IA affidabili e al graduale abbandono o rifiuto, da parte delle istituzioni pubbliche, di quelli considerati nocivi per la tenuta degli ordinamenti democratici. In questa prospettiva, sarà decisiva la capacità di intervento attribuita dalle legislazioni nazionali agli organismi indipendenti di controllo e l'esercizio effettivo, da parte di questi, di una potestà di vigilanza corrispondente, in grado di far emergere tutti i rischi derivanti dall'utilizzo delle applicazioni di IA suscettibili di incidere sulla protezione dei diritti e

delle libertà individuali, così da innescare un potenziale circolo virtuoso e favorirne l'impiego in modo responsabile.

D'altro canto, le segnalate limitazioni dell'ambito di applicazione della Convenzione potrebbero da sole compromettere il funzionamento della disciplina. Ciò vale sia per le esenzioni ratione materiae, che escludono i sistemi di IA operanti nei settori della sicurezza nazionale e della difesa dal controllo sull'osservanza degli obblighi convenzionali, sia per quelle ratione personae, che riservano agli Stati parti un ampio margine di discrezionalità circa l'applicabilità della Convenzione nei confronti di attori privati, sia, infine, per quelle relative alle attività di ricerca e sviluppo su sistemi di IA non ancora in uso.

Sommate tra loro, tali esenzioni incidono in misura rilevante sulla portata applicativa della Convenzione e, quel che più conta, sul perseguimento delle relative finalità. Esonerare dal controllo sul rispetto dei diritti umani e dei principi democratici i sistemi di IA utilizzati da attori privati, i quali, a loro volta, potrebbero essere coinvolti nello sviluppo di tecnologie applicabili al settore della sicurezza o a quello della difesa, determina una potenziale saldatura di interessi apertamente antitetici rispetto agli obiettivi che la Convenzione mira a conseguire, realizzando un compromesso al ribasso tra le ambizioni del regime di tutela promosso dall'accordo e la sua capacità di rappresentare un efficace strumento per garantire la legittimità democratica delle attuali e future applicazioni di IA.

Appare scontato, sotto questo profilo, operare un confronto con la disciplina normativa in vigore nell'Unione europea a partire dall'adozione dell'AI Act, che già vincola gran parte degli Stati membri del Consiglio d'Europa all'osservanza di regole puntuali in materia[48]. Com'è noto, tale strumento, accolto come un'iniziativa

[48] L'articolo 27, par. 2, della Convenzione di Vilnius prevede che gli Stati parti che siano membri dell'Unione europea "shall, in their mutual relations, ap-

pionieristica, è un atto legislativo di diretta applicazione negli ordinamenti nazionali, che ha introdotto un regime giuridico completo e dettagliato, volto a promuovere la diffusione di sistemi di IA affidabili nello spazio giuridico dell'Unione, non solo sotto i profili della sicurezza e della tutela dei diritti fondamentali, ma anche della capacità di operare nel mercato rilevante quali soggetti destinatari di finanziamenti o di altre forme di investimento legato allo sviluppo di tale tecnologia.

L'AI Act si rivolge, quindi, ai fornitori e ai distributori che operano nel mercato europeo, indipendentemente dalla loro provenienza geografica, adottando un "risk-based approach", in virtù del quale le applicazioni dell'IA sono classificate in base a diversi livelli di rischio, caratterizzati dall'osservanza di obblighi specifici[49]. Particolare enfasi è posta, altresì, sulla necessità di assicurare la trasparenza, imponendo la divulgazione dei contenuti, in particolare per le applicazioni di IA generativa.

Il regolamento dell'UE e la Convenzione di Vilnius presentano,

ply European Union rules governing the matters within the scope of this Convention without prejudice to the object and purpose of this Convention and without prejudice to its full application with other Parties. The same applies to other Parties to the extent that they are bound by such rules".

[49] Il livello di rischio più elevato, considerato inaccettabile, identifica i sistemi di IA che costituiscono una chiara minaccia per la sicurezza, l'assetto economico-sociale e la tutela dei diritti. Tali sistemi – tra cui possono includersi le applicazioni di IA preordinate a realizzare manipolazioni pregiudizievoli, talune forme di identificazione biometrica, o il cd. "social scoring" – sono esplicitamente vietati (articoli 6 ss. del regolamento). La categoria di sistemi considerati "ad alto rischio" comprende le applicazioni dell'IA suscettibili di arrecare danni significativi alla salute, alla sicurezza o ai diritti fondamentali delle persone. Si tratta di una categoria molto ampia, che ingloba i sistemi di IA operanti nelle infrastrutture critiche, nell'istruzione, nel mercato del lavoro, nell'accesso ai servizi essenziali, nell'identificazione biometrica a distanza, nell'applicazione della legge, nella gestione dei flussi migratori e del controllo delle frontiere e nell'amministrazione della giustizia (articoli 8 ss.).

pertanto, alcune sostanziali differenze – sia nell'ambito di applicazione (che nel caso della Convenzione è più ampio e copre le attività dell'intero ciclo di vita dei sistemi di IA potenzialmente interferenti con i diritti umani, la democrazia e lo Stato di diritto), sia nell'approccio normativo (più rigoroso e prescrittivo nel caso dell'IA Act, più generico e flessibile, perché basato sull'applicazione di principi fondamentali, quello della Convenzione), sia, infine, nelle loro finalità immediate (che l'IA Act individua nella regolamentazione del mercato dell'IA all'interno dell'Unione europea, allo scopo di garantire la sicurezza e proteggere i diritti fondamentali, ma in un quadro economico che incoraggi al contempo l'innovazione) – ma appaiono, al contempo, ispirati a logiche di regolamentazione complementari e suscettibili di mutuo apprezzamento.

Se, infatti, l'AI Act dimostra un forte potenziale di efficacia nell'ambito della sua giurisdizione, caratterizzandosi per un approccio selettivo basato su una dettagliata categorizzazione del rischio e sull'emarginazione delle applicazioni di IA con un elevato potenziale dannoso (che, al netto di alcune rigidità attuative, rafforza la fiducia degli utenti rendendo più accessibili le vie di ricorso), il rilievo attribuito dalla Convenzione di Vilnius alla valutazione del potenziale impatto dell'intero ciclo di vita di un sistema di IA sulla tutela dei diritti umani e dei valori democratici rappresenta una prospettiva di regolamentazione giudica più idonea contribuire alla creazione di una cultura globale dello sviluppo responsabile dell'IA, a patto che le deroghe previste dall'articolo 3 non finiscano per costituire un pretesto per svuotare di contenuto l'impianto normativo dell'accordo. Si può affermare, in questo senso, che nonostante le loro diverse finalità primarie, i due strumenti condividono i principali obiettivi di fondo ponendosi come modelli di riferimento per le iniziative successive.

Com'è stato correttamente osservato, è senz'altro possibile adottare una legislazione *ad hoc* che proibisca lo sviluppo di alcune tecnologie di IA[50] o il loro impiego in un particolare ambito, o che sottoponga il funzionamento dei sistemi di IA a restrizioni legali o a un regime di licenze o autorizzazioni, prevedendo opportune eccezioni per specifiche attività. Cosa diversa è introdurre misure in grado di prevenire l'impiego di una tecnologia di IA per finalità ostili o malevole: una sorta di "technical switch that would stop or turn off the AI technology if used for other objectives than specified in the business requirements"[51].

È rispetto a questo genere di interventi che sia il "risk-based approach" accolto nell'IA Act, sia i principi fondamentali incorporati nella Convenzione di Vilnius e il relativo sistema di garanzie e di controlli, saranno chiamati a dar prova di efficacia. L'utilità di questi regimi normativi, che costituiscono al momento l'avamposto sperimentale per una futura disciplina internazionale dei sistemi di IA, dovrà infatti misurarsi con la resistenza opposta dagli attori privati, giocoforza motivati pressoché esclusivamente da logiche di profitto nella loro attività di sviluppo e promozione di sistemi di IA sempre più sofisticati e autoreferenziali, e con la crescente fragilità dei sistemi democratici, messi costantemente in pericolo dalle aspirazioni di gruppi di potere di perpetuare indefinitamente la loro condizione di dominio sulla società.

Regole e controlli saranno determinanti per l'individuazione del punto di equilibrio tra l'inarrestabile progresso tecnologico e i suoi imprevedibili esiti e la salvaguardia della supremazia dei valori

[50] M. ŠMUCLEROVÁ, L. KRÁL, J. DRCHAL, *AI Life Cycle and Human Rights*, in A. QUINTAVALLA, J. TEMPERMAN, *Artificial Intelligence and Human Rights*, OUP, Oxford, 2023, 16 ss. Gli Autori citano l'esempio delle tecniche di manipolazione della coscienza di una persona, che potrebbero orientarne in senso antisociale il comportamento e condurre a forme di discriminazione (p. 34).

[51] *Ibidem.*

e delle istituzioni democratiche su qualsiasi diverso assetto organizzativo della società che sistemi di IA non trasparenti potrebbero contribuire, in modo decisivo, a realizzare. L'alternativa è una "resa fiduciosa" al primato nichilista della tecnica, che reca con sé il rischio di favorire una nuova deriva autoritaria travestita da paternalismo indulgente e legittimata dall'adozione di inconsistenti norme-compendio, aventi una mera funzione descrittiva.

Al contrario, la sfida cruciale per il futuro della regolamentazione dell'IA risiede nella capacità della comunità internazionale di superare un simile approccio, che si limita a inseguire il progresso tecnologico con interventi normativi frammentari e spesso inadeguati, e di ancorare saldamente il regime di legittimazione dell'impiego dell'IA al rispetto di principi universali. L'istituzione di un "guardiano dell'IA" universale, sotto forma di un organismo internazionale specializzato, dotato di adeguate competenze e meccanismi di monitoraggio e accountability, potrebbe rappresentare un passo fondamentale in questa direzione, rafforzando la fiducia nell'ordine giuridico internazionale ed evitando che una tecnologia dalle potenzialità straordinarie si trasformi in uno strumento di compressione delle libertà e di esacerbazione delle disuguaglianze a livello globale in questo momento storico cruciale.

AI nel rapporto di cura

di *Candida Conti*[*]

SOMMARIO: 1. Premessa. – 2. Il paziente nella società dell'informazione. – 2.1. Il paziente nella società dell'algoritmo. – 3. Il medico nella società dell'algoritmo. – 4. Riflessioni conclusive.

1. Premessa

Il rapporto tra medico e paziente si fonda sul principio dell'autonomia reciproca, in base alla quale entrambe le parti conservano il diritto all'autodeterminazione. La libertà di scelta rappresenta un pilastro essenziale del rapporto di cura: il paziente deve poter ricevere informazioni complete e comprendere pienamente i trattamenti proposti, mentre il professionista sanitario deve poter esercitare il proprio giudizio clinico liberamente e con consapevolezza[1], basandosi sulle proprie competenze scientifiche e sull'esperienza.

Questa concezione del rapporto di cura trova un solido fondamento nell'articolo 32 della Costituzione italiana, che tutela il diritto alla salute non solo come interesse collettivo ma anche come libertà individuale di partecipare attivamente al proprio percorso terapeutico che può essere limitato esclusivamente in casi eccezionali,

[*] Dottoranda in Società in mutamento: politiche, diritti e sicurezza, Università degli Studi della Tuscia.

[1] La Corte cost. nella sentenza 282/2002 precisa che il medico "opera le scelte professionali basandosi sullo stato delle conoscenze a disposizione".

espressamente previsti dalla legge e, in ogni caso, nel rispetto della persona umana[2].

Prima del codice di Norimberga[3] difficilmente veniva messa in discussione la scelta dell'operatore sanitario, ritenuta in ogni caso orientata al bene del soggetto sottoposto a trattamento. Il modello dominante era quello paternalistico, nel quale il medico agiva come unico decisore e il paziente si limitava ad assecondare le indicazioni, affidandosi integralmente al suo sapere.

Solo in un momento successivo si è iniziato a riconoscere maggiore rilevanza alla volontà del singolo nel proprio percorso di cura[4], prendendo consapevolezza della possibilità che anche il medico

[2] Il comma 2 dell'art. 32 dice che nessuno può essere obbligato ad un determinato trattamento sanitario: il principio contenuto in tale disposizione comporta che il cittadino può legittimamente reagire contro ogni intervento dell'autorità pubblica. Il cittadino è libero sottoporsi o meno ad un intervento sanitario, salvo il caso non vi sia un'obbligatorietà per legge diversamente può solo limitarsi a reprimere gli abusi del sanitario quando essi costituiscono il fatto colposo. Giovanni de Cesare, voce Sanità (dir. amm) in Enciclopedia del diritto, XLI Rive -Sepa, 246).

[3] Il Codice di Norimberga del 1947 si apre dicendo "Il consenso volontario del soggetto umano è assolutamente essenziale". "nessuno può essere obbligato ad un determinato trattamento sanitario" bisogna riconoscere un diritto all'autodeterminazione in ordine al rifiuto della cura come una libertà individuale e come al "limite" al potere pubblico, in quanto è affidato alla legge il compito di valutare l'operatività degli altri poteri dello Stato; tale decisione in Assemblea Costituente è dovuta al ricordo della scoperta degli esperimenti sugli esseri umani nei campi di concentramento nazisti come ricordato da S. MANGIAMELI, *Autodeterminazione: diritto di spessore costituzionale?*, in «Forum dei Quaderni costituzionali, Relazione tenuta al IV Laboratorio Sublacense su «La comunità familiare e le scelte di fine vita», 3-5 luglio 2009, 18.

[4] La legge 833/1978 considera la volontà del paziente quale elemento fondante delle scelte terapeutiche, ugualmente la Convenzione di Oviedo del 1997 all'art. 5 considera il consenso informato come una necessità per l'avvio dei trattamenti terapeutici individuati dal medico, seppur revocabile; la Carta dei diritti fondamentali dell'Unione europea all'art. 3 considera il consenso informato come la ba-

possa errare e, di conseguenza, della necessità di instaurare un dialogo autentico con la persona interessata. Da qui l'emersione progressiva del consenso informato come strumento di valorizzazione degli interessi del paziente[5] per il raggiungimento dell'obiettivo di benessere psico-fisico in una prospettiva di cooperazione.

Nonostante la maggiore centralità attribuita al paziente, quest'ultima resta in ogni caso condizionata dal persistente squilibrio conoscitivo e per questa ragione, dottrina e giurisprudenza si sono interrogate a lungo sull'effettiva libertà del consenso, sottolineando come esso possa dirsi realmente informato solo a condizione che il paziente possa ricevere informazioni complete, esaustive e formulate in modo comprensibile, tenendo conto anche del suo livello socioculturale[6].

se giuridica necessaria per riconoscere la legittimità dei trattamenti sanitari nella relazione tra medico e paziente; sulla libertà di scelta e sull'individuazione e condivisione delle rispettive autonomie e responsabilità cfr. Codice di deontologia medica al titolo III, art. 20 ss.. L'art. 35, in particolare, impone al medico di astenersi dal trattamento sanitario qualora questo venga rifiutato dal paziente interessato.

[5] Specificato nell'art. 33 c.1 della L. 833/1978 "Gli accertamenti ed i trattamenti sanitari sono di norma volontari. Nei casi di cui alla presente legge e in quelli espressamente previsti da leggi dello Stato possono essere disposti dall'autorità sanitaria accertamenti e trattamenti sanitari obbligatori, secondo l'art. 32 della Costituzione, nel rispetto della dignità della persona e dei diritti civili e politici, compreso per quanto possibile il diritto alla libera scelta del medico e del luogo di cura". Sull'importanza di autodeterminazione nel percorso di cura "*Voluntas aegroti suprema lex*" *cfr. ex multis* N.VICECONTE, *La sospensione delle terapie salvavita: rifiuto delle cure o eutanasia? Riflessioni su autodeterminazione e diritto alla vita tra testo costituzionale e applicazioni giurisprudenziali*, in «Diritto e Società», (Relazione al Convegno "Il testamento biologico e la fine della vita, organizzato dall'OMCeO della Provincia dell'Aquila, L'Aquila, 23 ottobre 2010), 165-174.

[6] Nell'art. 1 comma 3 della Legge 219/2017 si scrive che la persona deve essere informata «in modo completo e aggiornato e a lei comprensibile riguardo alla diagnosi, alla prognosi, ai benefici e ai rischi degli accertamenti diagnostici e dei trattamenti sanitari indicati, nonché riguardo alle possibili alternative e alle conse-

In questo contributo concentreremo l'attenzione sul consenso espresso dal paziente come garanzia di un rapporto terapeutico condiviso, analizzeremo in particolare le trasformazioni che stanno interessando tale rapporto: prima nella società dell'informazione, in cui l'accesso alle informazioni ha ampliato il ruolo attivo del paziente, poi nella società degli algoritmi in cui emergono difficoltà specifiche del rapporto tradizionalmente conosciuto. In quest'ambito l'operatore sanitario è chiamato a spiegare processi decisionali generati da sistemi di intelligenza artificiale che, allo stato attuale della conoscenza tecnologica, non può comprendere pienamente, suscitando interrogativi sulla tenuta del paradigma di cura.

2. Il paziente nella società dell'informazione

Nella società dell'informazione le persone possono accedere a una pluralità di nuovi canali informativi, in particolare digitali, che, almeno in astratto, riducono il tradizionale divario conoscitivo esi-

guenze dell'eventuale rifiuto del trattamento sanitario e dell'accertamento diagnostico o della rinuncia ai medesimi…»; Già prima del legislatore del 2017 la Corte di cassazione si era pronunciata in tal senso con sentenza del 9 febbraio 2010, n. 2847. L'obbligo informativo può arrestarsi solo di fronte «ai soli rischi imprevedibili, ovvero agli esiti anomali, al limite del fortuito», ma «al di là di tale limite, il professionista sanitario ha l'obbligo di fornire al paziente, in modo […] dettagliato, tutte le informazioni scientificamente possibili sull'intervento chirurgico che intende eseguire, sulle conseguenze normalmente possibili sia pure infrequenti (*id quod plerumque accidit*): così spiegato dalla Suprema Corte di cassazione con sentenza n. Cass., sent. 11 dicembre 2013, n. 27751; ugualmente la Corte Cost. con sentenza 15 dicembre 2008, n. 438, chiarisce che le informazioni "devono essere le più esaurenti possibili, proprio al fine di garantire la libera e consapevole scelta da parte del paziente". Sul tema *cfr. ex multis* N. POSTERARO, *Osservazioni sul consenso informato alla luce di giurisprudenza e dottrina recenti: dai profili di responsabilità civile e penale del sanitario alla spersonalizzazione del rapporto medico-paziente*, in «Amministrazione in Cammino», 2014, 6 ss.

stente rispetto al professionista della salute. L'ampia accessibilità alle informazioni sanitarie, unita alla crescente diffusione di strumenti tecnologici, contribuisce a ridefinire il ruolo del paziente, non più mero destinatario passivo delle decisioni cliniche, ma come soggetto maggiormente partecipativo.

Le persone sono immerse in un flusso informativo continuo e crescente, alimentato non solo da fonti *online*, spesso eterogenee, ma anche da dispositivi di *m-health* che consentono il monitoraggio costante del proprio stato di salute. Strumenti come *smartwatch* o *smart ring* permettono di rilevare battiti cardiaci, ossigeno del sangue, qualità del sonno o parametri del ciclo mestruale, mentre dispositivi più specifici, come le *smart pen,* registrano addirittura i livelli di insulina.

Questa evoluzione sembra rappresentare una prospettiva positiva per l'alleanza terapeutica, coerente con il principio di autodeterminazione sancito dall'art. 32 della Costituzione. Tuttavia, non possono essere trascurati i profili critici che emergono in concreto: sempre più spesso, il paziente si presenta alla visita con ipotesi diagnostiche e terapeutiche elaborate autonomamente attraverso fonti *online*, che non sono sempre affidabili né correttamente interpretate. Ne consegue che il confronto con il medico possa assumere toni conflittuali, orientandosi più verso la difesa di pretese che verso un percorso di cura condiviso.

Inoltre, questi strumenti espongono l'assistito al rischio di sovraccarico informativo, che può compromettere la capacità di selezionare, comprendere e valutare criticamente le informazioni effettivamente rilevanti ai fini delle scelte cliniche e terapeutiche. In altri termini, l'accesso ampio alle informazioni, se da un lato favorisce l'autonomia del paziente, dall'altro può generare confusione e tensioni nel rapporto di cura rendendo più complesso il raggiungimento di decisioni condivise.

Proprio per questo, l'obbligo del medico di informare non può ritenersi superfluo o ridimensionato; al contrario, l'eccesso di informazioni rende il ruolo del professionista ancora più centrale, come garante della correttezza, pertinenza e comprensibilità delle informazioni trasmesse. In questo contesto, il professionista deve aiutare il soggetto sottoposto alle cure a distinguere tra informazioni utili e dati fuorvianti, orientando la scelta valutativa su basi scientificamente fondate.

Il consenso informato sembra non perdere la sua funzione, ma assume, piuttosto, una valenza ancora più significativa: diventa lo strumento di mediazione tra la conoscenza tecnica e l'autodeterminazione del paziente. L'informazione necessaria a esprimere un consenso si può ritenere adeguata quando il professionista ha illustrato chiaramente i rischi e i benefici del trattamento, o del "non trattamento", consentendo al soggetto sottoposto a diagnosi di esercitare una scelta consapevole.

Per questi motivi, appare giuridicamente improprio mettere in discussione l'utilità del consenso informato nell'ambito del trattamento sanitario: anche nella società dell'informazione, esso resta il mezzo principale attraverso cui si esercita il diritto fondamentale alla salute[7].

2.1. Il paziente nella società dell'algoritmo

Alla luce di quanto osservato in merito alla funzione del consenso informato nella società dell'informazione, è necessario spostare lo sguardo verso la società dell'algoritmo. In questo contesto, il consenso si confronta con sfide nuove, legate alla natura opaca dei processi decisionali automatizzati, che risultano spesso difficili

[7] R. BROWNSWORD, *Il consenso informato nella società dell'informazione*, in «Salute e Società», 3, 2012, 171-178.

da interpretare. Tale complessità rende arduo per il medico fornire all'assistito informazioni chiare e comprensibili sui rischi e benefici del trattamento, sollevando interrogativi sulla possibilità di garantire un consenso pienamente libero e consapevole[8].

Se è ormai consolidato che il medico che utilizza strumenti digitali o sistemi di intelligenza artificiale debba informare il paziente del loro impiego[9], resta invece aperta la questione di come rendere effettivamente comprensibili i vantaggi e gli svantaggi derivanti dall'uso di tali tecnologie[10].

L'adozione dell'intelligenza artificiale rischia di accentuare, dunque, l'asimmetria informativa già presente nel rapporto terapeutico, poiché i processi algoritmici rimangono, almeno allo stato attuale, poco trasparenti, nonostante i progressi nell'ambito dell'*explainable artificial intelligence.*

Oltre all'opacità dei sistemi, occorre considerare la loro capacità persuasiva, che può influenzare in modo particolare soggetti vul-

[8] D. MORANA, T. BALDUZZI, F. MORGANTI, *La salute "intelligente": eHealth, consenso informato e principio di non discriminazione*, in «Federalismi.it», 34, 2022, 139.

[9] Come chiarito dal Comitato Nazionale per la Biosicurezza e le Biotecnologie le scienze della vita e il Consiglio Superiore di Sanità provano ad arginare a monte il problema stabilendo un obbligo etico e giuridico di informare coloro che si sottopongono a trattamenti sanitari così innovativi relativamente alle modalità di applicazione di strumenti potenziati con AI per la diagnosi e la terapia e in quali ambiti e limiti è previsto il controllo umano della macchina, in COMITATO NAZIONALE PER LA BIOETICA E COMITATO NAZIONALE PER LA BIOSICUREZZA, LE BIOTECNOLOGIE E LE SCIENZE DELLA VITA, I*ntelligenza artificiale e medicina: aspetti etici*, 29 maggio 2020, 14.

[10] P. ZUDDAS, *Intelligenza artificiale in medicina: alcune risposte – significative ma parziali – offerte dal codice di deontologia medica (in materia di non discriminazione, consenso informato e relazione di cura)*, in «Rivista italiana di informatica e diritto», 2, 2024, 17 scrive della mancanza di un chiaro riferimento normativo che indichi la quantità e la qualità delle informazioni che sono da comunicare al paziente.

nerabili, già in condizioni di bisogno legate a problemi di salute. Studi pubblicati su riviste scientifiche evidenziano come l'uso intensivo di modelli di linguaggio di grandi dimensioni (*Large Language Models*) possa generare forme di dipendenza psicologica, favorendo atteggiamenti di *overdependence* – ossia una fiducia eccessiva nello strumento tecnologico – con possibili ricadute in termini di ansia, stress e depressione. Questo fenomeno può portare a un ampliamento dei bisogni di cura e assistenza e avere un impatto sulla percezione sociale della funzione medica, sminuendo l'autorevolezza e il ruolo del medico nel percorso clinico e terapeutico.

Alcuni esponenti dell'industria tecnologica hanno apertamente dichiarato di attendersi che l'intelligenza artificiale possa assumere una funzione di "super-persuasione", ovvero la capacità di influenzare e orientare opinioni, desideri e bisogni degli individui con un'efficacia superiore a qualsiasi forma tradizionale di persuasione umana. Mentre in passato aziende come Purdue Pharma hanno dovuto investire ingenti risorse economiche per orientare le scelte prescrittive dei medici, oggi sistemi algoritmici sofisticati possono ottenere risultati analoghi con costi significativamente inferiori[11].

In questo scenario, il rischio è che il paziente, di fronte a risposte algoritmiche apparentemente plausibili e coerenti, sia indotto a riporre una fiducia eccessiva nei sistemi di intelligenza artificiale, fino a pratiche di autodiagnosi o di gestione terapeutica autonoma, entrambe potenzialmente pericolose. Il passaggio dal "dottor Google" al "dottor ChatGPT" segna, infatti, un salto qualitativo rilevante: non si tratta più di semplici interrogazioni generiche, ma di una profonda interazione tra dati biologici e modelli probabilistici

[11] P. BENANTI, *L'ontologia del rischio: dal farmaco all'algoritmo. La fallacia dell'autoregolamentazione del mercato di fronte alla vulnerabilità umana*, 18 dicembre 2025.

di linguaggio, capaci di fornire risposte verosimili anche senza una reale comprensione medica del caso concreto[12].

Tuttavia, al di là dei rischi legati alla persuasione algoritmica e all'eccessiva fiducia nei sistemi di intelligenza artificiale, occorre sottolineare un limite intrinseco di tali strumenti: l'assenza di competenza empatica. L'empatia rappresenta un elemento centrale del rapporto di cura e proprio la sua mancanza negli algoritmi avvalora l'importanza insostituibile del medico. Solo il professionista umano può comprendere e gestire dimensioni profondamente umane, quali la paura di ammalarsi, il dolore, il timore di menomazioni o della morte, fattori che influenzano in modo determinante le scelte terapeutiche e l'adesione al percorso[13].

Si pensi, ad esempio, al caso di una giovane donna a cui venga diagnosticato un tumore al seno: nella valutazione delle diverse opzioni terapeutiche che possono precedere una mastectomia, il medico è chiamato a bilanciare non solo i dati clinici, ma anche fattori extra-medici quali: età, storia personale e familiare, vissuto emotivo, percezione della propria identità corporea e aspettative di vita. Questi elementi, essenziali per un percorso di cura personalizzato e condiviso, difficilmente possono essere adeguatamente colti e integrati da un sistema algoritmico, che resta privo della sensibilità e della capacità relazionale proprie dell'essere umano.

Nonostante l'aumento dell'utilizzo di piattaforme digitali in grado di fornire consulenze ovunque e in qualsiasi momento – per questo apprezzati dai cittadini, soprattutto più giovani – resta evidente l'importanza insostituibile dell'empatia nel rapporto di cura. Il caso del *Symptom Checker Chatbot*, sviluppato dalla società lon-

[12] *Ibidem*.

[13] J. P. Solè, *Il regolamento dell'unione europea sull'intelligenza artificiale, la discrezionalità amministrativa e la riserva di umanità*, in «Rivista Trimestrale di Diritto Pubblico», 3, 2024, 835-839.

dinese *Babylon*, fondata nel 2013 e operante in convenzione con il sistema sanitario britannico, lo dimostra chiaramente. Il *software*, progettato per fornire una prima valutazione clinica e, se necessario, indirizzare l'utente verso l'operatore di medicina generale, non ha avuto il successo atteso: gli utenti tendevano comunque a richiedere successivamente un consulto tramite videochiamata o di persona[14].

Questo esempio evidenzia come, sebbene gli strumenti digitali possano rappresentare un valido supporto, nessuna macchina può sostituire la funzione "umana" del medico, capace di offrire comprensione, ascolto e mediazione empatica, elementi fondamentali per garantire un percorso di cura efficace e condiviso.

3. Il medico nella società dell'algoritmo

Nella scelta clinica o terapeutica, il medico si trova oggi non più di fronte a un semplice bivio, ma a un vero e proprio crocevia di opportunità e rischi. Alla tradizionale tensione tra il rispetto delle evidenze scientifiche e l'esercizio del giudizio clinico autonomo fondato sulla peculiarità del caso concreto si aggiunge una terza opzione: il ricorso ai suggerimenti forniti da sistemi di intelligenza artificiale.

Nei primi due casi, il professionista beneficia di un assetto di responsabilità relativamente definito dalla normativa vigente, potendo dimostrare di aver rispettato le linee guida o di essersene consapevolmente discostato in ragione delle specificità cliniche del paziente. Diversa è invece la situazione quando il medico decide di

[14] A. Melva, V. Zurlo, *Babylon Health: il servizio anglosassone di medicina generale privato basato su app*, in «Rivista Società italiana di Medicina Generale», 5, 2019, 15-17.

avvalersi di sistemi di intelligenza artificiale, soprattutto di quelli basati su algoritmi di *machine learning*, caratterizzati da opacità decisionale e da un elevato grado di autonomia. In tale ipotesi, il professionista è consapevole di non poter ricostruire integralmente il percorso logico che conduce all'*output* della macchina, con conseguente difficoltà di controllo e di spiegazione del risultato[15].

Questa opacità alimenta una diffidenza del personale sanitario nei confronti delle nuove tecnologie, anche e soprattutto in ragione dell'incertezza sui confini della loro responsabilità giuridica. Il rischio è quello di un'espansione della medicina difensiva: il medico potrebbe decidere di sottoporre comunque il paziente a ulteriori esami e accertamenti, al solo fine di ridurre il rischio di contenzioso, oppure scegliere di non utilizzare affatto il sistema intelligente, salvo che esso sia pienamente certificato o espressamente autorizzato dall'autorità pubblica.

Un'ulteriore, seppur più complessa, strategia potrebbe consistere nel tentativo di compensare l'opacità algoritmica mediante una documentazione clinica particolarmente accurata e strutturata. Tale approccio, in linea con i principi della legge n. 24 del 2017, c.d. legge Gelli-Bianco, impone tuttavia al medico di dimostrare *ex post* di aver esercitato un autonomo giudizio professionale anche in presenza di *output* automatizzati. Ciò presuppone, oltre a solide competenze cliniche, anche adeguate competenze tecnico-digitali, che spesso il personale sanitario non possiede e difficilmente potrà acquisire senza investimenti in programmi di formazione specifica.

L'importanza della formazione emerge chiaramente da un recente caso verificatosi nei Paesi Baschi, dove è stato sperimentato il sistema *Quantus Skin* per la valutazione del rischio di tumore cuta-

[15] Al Principio 13 della Carta europea di etica medica si scrive: "Il medico deve informare il paziente se le condizioni morali e tecniche sono tali da impedirgli di agire in piena autonomia".

neo e la definizione delle priorità di accesso alle visite dermatologiche. Il fallimento della sperimentazione non è dipeso tanto da un malfunzionamento dell'algoritmo, quanto dall'incapacità dei medici di medicina generale di fornire *input* adeguati al sistema. Ne sono derivate valutazioni diagnostiche errate, con conseguenze gravi per la salute dei cittadini, soprattutto considerando la necessaria tempestività nella diagnosi del melanoma[16].

Alla luce di tali criticità, si impone una riflessione sulla responsabilità civile del professionista sanitario che utilizza strumenti basati sull'intelligenza artificiale. Al riguardo, parte della dottrina ipotizza una responsabilità oggettiva o semi-oggettiva *ex* art. 2050 c.c., qualificando l'attività sanitaria, assistita da sistemi di AI, come attività pericolosa. Tale impostazione è stata contestata da chi ritiene forzata l'equiparazione di attività pericolosa con l'attività sanitaria svolta attraverso strumenti intelligenti; pertanto, altri ritengono più appropriato il richiamo all'art. 2051 c.c., relativo al danno cagionato da cose in custodia, ma anche questa soluzione presenta evidenti limiti, poiché presuppone un controllo umano spesso assente nei sistemi algoritmici opachi e auto-apprendenti, perché come insegna il dibattito sviluppatosi a partire dal test di Turing, all'aumentare dell'intelligenza della macchina diminuisce la sua comprensibilità[17].

Sembra evidente che questo percorso evolutivo richieda un nuovo assetto regolatorio che sappia tenere il passo con l'evoluzione tecnologica, consapevoli del fatto che non possiamo ricor-

[16] A. BERNARDO, M. ALVAREZ DEL VAYO, C. TORRECILLAS, A. MAQUEDA, *Lunar or cancer?, El algoritmo que se equivoca en uno de cada tres melanomas y obvia a los pacientes con la piel oscura*, in «Civio», 26 giugno 2025, spiegano che il sistema avrebbe ignorato circa un melanoma su tre, soprattutto di pazienti con la pelle scura, producendo evidenti effetti discriminatori nell'accesso alle cure.

[17] T. CANTELMI, L. LODEVOLE, *Nuove prassi e norme per un sistema sanitario "smart": ricadute dell'AI Act sulla relazione di cura*, in «Rivista Italiana di Informatica e Diritto», 2, 2024, 4.

rere all'analogia con le categorie tradizionali, anche perché c'è una pluralità di soggetti coinvolti e potenzialmente responsabili oltre il medico: creatore del *software*, produttore, fornitore e da ultimo anche l'utente.

Il legislatore europeo ha chiarito che le norme nazionali vigenti non sono adeguate a considerare la responsabilità del professionista sanitario, che si avvale di tali strumenti e per tale ragione prefigura la proposta di una direttiva sull'adeguamento delle norme in materia di responsabilità civile extracontrattuale, introducendo strumenti di accesso ai dati rilevanti e prevede presunzioni di colpa in presenza di particolari difficoltà probatorie[18].

L'efficacia di tali strumenti dipenderà dalla capacità dell'interprete di ricostruire il nesso causale tra la condotta e l'evento dannoso, distinguendo tra ipotesi di malfunzionamento del sistema e casi di insufficiente o mancato controllo umano.

[18] Come spiegato nella proposta di direttiva del Parlamento europeo e del Consiglio relativa all'adeguamento delle norme in materia di responsabilità civile extracontrattuale all'intelligenza artificiale (direttiva sulla responsabilità da intelligenza artificiale) 2022/0303 (COD) COM(2022) 496 final del 28 settembre 2022 all'art. 4 comma 6, per cui in "caso di sistemi di IA che non sono ad alto rischio, l'articolo 4, paragrafo 5, stabilisce una condizione per l'applicabilità della presunzione di causalità, in base alla quale quest'ultima è soggetta alla condizione che l'organo giurisdizionale stabilisca che è eccessivamente difficile per l'attore dimostrare l'esistenza del nesso causale. Tali difficoltà devono essere valutate alla luce delle caratteristiche di determinati sistemi di IA, come l'autonomia e l'opacità, che nella pratica rendono estremamente difficile la spiegazione del funzionamento interno del sistema di IA e pregiudicano la capacità dell'attore di dimostrare l'esistenza del nesso di causalità tra la colpa del convenuto e l'output del sistema di IA". Tale costatazione è inserita nell'ambito di una più generale consapevolezza della difficoltà di stabilire un nesso causale tra la non conformità e l'output prodotto dal sistema di IA o la mancata produzione di un output da parte del sistema di IA che ha cagionato il danno. La presunzione è volta a garantire un equo risarcimento del danneggiato.

In questa fase in mancanza di un assetto chiaro delle responsabilità, il rischio è quello del paradosso per cui il medico, per timore di conseguenze giuridiche, potrebbe rinunciare all'uso di strumenti diagnostici innovativi, privando l'assistito della possibilità di accedere a cure potenzialmente migliori. Ciò finirebbe per compromettere non solo l'autodeterminazione professionale del sanitario, ma anche l'effettiva tutela del diritto alla salute.

4. Riflessioni conclusive

Da questa analisi emerge come si stia assistendo a un'evoluzione del paradigma di cura, determinata in larga misura da una crescente consapevolezza del paziente, favorita dalla disponibilità di molteplici canali informativi e strumenti digitali. Tale evoluzione si innesta, tuttavia, in un contesto caratterizzato da una progressiva perdita di fiducia nel sistema sanitario, che rischia di essere ulteriormente aggravata dall'introduzione pervasiva delle tecnologie digitali e dei sistemi di intelligenza artificiale in grado di dare all'assistito la percezione di poter fare da sé. In questo scenario, non viene messo in discussione soltanto il rapporto terapeutico nella sua configurazione tradizionale, ma si assiste alla creazione di un modello in cui il medico si trova a inseguire nuovi bisogni di cura e di assistenza, vedendo progressivamente svuotata la propria libertà di analisi clinica a causa del timore di contenziosi giudiziari che inducono il professionista a ridimensionare il proprio operato, privilegiando soluzioni percepite come più "sicure". Il consenso informato rischia di perdere la propria funzione sostanziale di strumento di partecipazione consapevole, per ridursi a un mero meccanismo di medicina difensiva.

Di fronte a questa evoluzione problematica del paradigma, si è

inteso dimostrare la necessità del ruolo del professionista della salute all'interno del rapporto terapeutico, dimostrando come la scienza medica non possa essere ridotta a una semplice elaborazione di dati, ma deve essere intesa come un'attività ermeneutica, fondata sull'empatia, sull'intuizione clinica e sulla comprensione del contesto biografico del soggetto sottoposto alle cure. Le capacità umane rappresentano un elemento imprescindibile del processo terapeutico, che nessun algoritmo, per quanto sofisticato, è in grado di replicare. La crescente interposizione di interfacce conversazionali e sistemi automatizzati rischia di ridurre tutto ad una duplice semplificazione: il paziente viene ridotto a una sommatoria di parametri e il medico progressivamente trasformato in un mero validatore di *output* tecnologici, con un conseguente svuotamento dell'anamnesi e della relazione terapeutica del loro significato umano e professionale.

Per tali ragioni, appare necessario orientarsi verso la costruzione di un modello integrato, capace di coniugare il meglio dell'innovazione tecnologica con il valore insostituibile dell'empatia umana. Una sostituzione integrale del professionista sanitario, infatti, risulta non solo lontana, ma anche incompatibile con l'essenza stessa del paradigma di cura. Da ciò discende l'urgenza di un intervento legislativo volto a delineare con precisione il regime delle responsabilità connesse all'impiego dell'intelligenza artificiale in sanità, al fine di preservare la fiducia nel sistema e, al contempo, consentire al professionista di adottare l'innovazione senza timori ingiustificati. Si auspica un adeguamento normativo che eviti tanto l'attribuzione indiscriminata della colpa quanto il rischio opposto di zone di impunità, al fine di garantire al medico di poter operare in equilibrio tra l'intelligenza umana e l'intelligenza artificiale.

La *Constitutional AI* di Anthropic tra normatività, legittimazione e limiti

di Giovanni D'Alessandro

SOMMARIO: 1. Prologo. La *Constitutional AI* come *Grenzfall* del costituzionalismo. – 2. Le tre matrici del costituzionalismo. – 3. La CAI come sistema normativo autoreferenziale. – 4. Tre deficit e una tesi intermedia. – 5. La privatizzazione del potere tecnico e il paradosso della legittimazione retorica. – 6. I contro-modelli istituzionali. AI Act e legge italiana n. 132 del 2025. – 7. Epilogo. La risposta tecnica e la risposta costituzionale.

1. Prologo. La *Constitutional AI* come *Grenzfall* del costituzionalismo

Il presente saggio esamina la *Constitutional AI* (d'ora in avanti CAI) di Anthropic come caso-limite (*Grenzfall*) del costituzionalismo nell'età algoritmica. Più che un oggetto tecnologico da descrivere, è un reagente teorico che, scrutinato con gli strumenti del diritto costituzionale comparato, rivela per contrasto le condizioni necessarie di ogni legittimo esercizio del potere normativo che incida con effetti generali su posizioni fondamentali e aspiri a tenere luogo delle decisioni pubbliche.

Un'impresa privata ha denominato "costituzione" l'insieme di princìpi che incorpora nel *training* dei propri modelli linguistici di grandi dimensioni[1], li ha ispirati alla Dichiarazione universale dei

[1] Y. BAI et al., *Constitutional AI: Harmlessness from AI Feedback*, in *ArXiv.org* (15 dicembre 2022); ANTHROPIC, *Claude's Constitution* (2026), in *An-*

diritti dell'uomo del 1948, ha condotto esperimenti di partecipazione pubblica alla loro redazione e ha costruito una gerarchia normativa analoga, sul piano della forma, a quella degli ordinamenti giuridici. Questo contributo accetta la sfida implicita in queste scelte lessicali. E intende verificarne le implicazioni logiche fino alle loro conseguenze costituzionali.

Conviene perciò muovere dalla parola prima che dalla cosa, chiedendo non che cosa sia una costituzione ma che cosa faccia, in ciascun uso, chi la nomina. Gli usi che Anthropic ne ha fatto sono quattro e non coincidono.

Nel 2022 il termine è scelto per metonimia dal metodo, perché il sistema si lascia addestrare «entirely through the specification of a short list of principles or instructions, i.e. a constitution», e costituzione designa allora una lista di istruzioni rivolte a un processo di addestramento, mentre "costituzionale" qualifica una tecnica, non un documento. Nel 2023 la parola designa un parametro di giudizio, giacché «the system uses a set of principles to make

thropic.com/constitution (21 gennaio 2026), rilasciata sotto licenza CC0 1.0. Per l'ispirazione dei princìpi originari alla Dichiarazione universale dei diritti dell'uomo (oltre che alle *Sparrow Rules* di DeepMind e ai *Terms of Service* di Apple) e per gli esperimenti di partecipazione pubblica, v. ANTHROPIC, *Claude's Constitution* (2023), in *Anthropic.com/news/claudes-constitution* (9 maggio 2023), nonché *infra* nt. 5. Per i quattro usi del termine censiti nel testo: BAI et al., cit., § 1 («We chose the term 'constitutional' because we are able to train less harmful systems entirely through the specification of a short list of principles or instructions, i.e. a constitution») e § 1.1 («a simple list of natural language instructions or principles»); ANTHROPIC, *Claude's Constitution* (2023), cit., per la frase sul parametro di giudizio; ANTHROPIC, *Claude's Constitution* (2026), cit., apertura, per la descrizione delle intenzioni e l'autorità finale, sezione *Concluding thoughts*, § *On the word "constitution"*, per la glossa sul termine, e sezione *Avoiding problematic concentrations of power* per l'uso politico dell'aggettivo («Circumventing constitutional limits on power», «seize or retain power in an unconstitutional way»). Il documento, che si dichiara «a perpetual work in progress», è citato nella versione del 21 gennaio 2026 (ed. webPDF 26-02.02a), consultata il 20 ottobre 2026.

judgments about outputs, hence the term "Constitutional"», e la funzione evocata è quella della norma-parametro, esercitata però dal sistema su sé stesso nell'addestramento. Nel 2026 il documento si presenta come «a detailed description of Anthropic's intentions for Claude's values and behavior» e insieme come «the final authority on our vision for Claude», cioè come dichiarazione di intenti munita di supremazia, e nelle considerazioni conclusive glossa infine il proprio titolo, avvertendo che il senso cercato è «closer to what "constitutes" Claude», come «a person's constitution is their fundamental nature and composition», sicché una costituzione in questo senso «is less like a cage and more like a trellis».

Lista di istruzioni, parametro, dichiarazione sovraordinata di volontà, complessione. I quattro usi condividono un nucleo, un insieme di princìpi che governa il comportamento del sistema, ma nessuno contiene il tratto che la parola porta con sé dal 1789, la limitazione del potere di chi la pone, e l'ultimo se ne allontana nel momento in cui rifiuta l'immagine della gabbia, salvi i limiti che il testo impone alla stessa Anthropic e che restano, senza garanzia istituzionale indipendente, nella disponibilità di chi li ha scritti. Quel tratto il documento lo conosce bene, perché lo impiega nel senso ordinario quando vieta al modello di concorrere a «circumventing constitutional limits on power», e il senso garantista dell'aggettivo resta così riferito al potere politico, anche a quello che Anthropic volesse esercitare per il tramite di Claude, mentre per sé, cioè per il testo che governa il modello, sceglie quello fisiologico, la costituzione come complessione di un individuo.

In queste pagine 'costituzionale' è usato sempre nel senso garantista. Gli altri, dal metonimico del 2022 al fisiologico del 2026, vi compaiono soltanto per riferire l'uso altrui. La tesi che s'intende svolgere è la seguente. La CAI è il più sistematico, fra quelli esaminati, dei tentativi privatistici di trasposizione del lessico costituzio-

nale nel dominio dell'allineamento algoritmico, ma questa trasposizione rivela che il costituzionalismo non è riducibile a una grammatica interna di sicurezza, perché richiede, per le scelte che incidono con effetti generali su posizioni fondamentali e pretendono di tenere luogo delle decisioni pubbliche, condizioni esterne irriducibili (pubblicità delle fonti prodotte da procedure di legittimazione democratica, controllabilità esterna affidata a organi indipendenti, contestabilità procedurale giuridicamente azionabile, responsabilità verso la comunità politica) che nessun sistema di *training* può produrre dall'interno. Il criterio adottato non è arbitrario. È Anthropic, evocando il linguaggio del costituzionalismo, ad aver scelto il campo da gioco.

La tesi non implica che la CAI sia normativamente insignificante. Essa possiede un valore d'uso normativo reale. Orienta il comportamento del modello e rende pubblici i criteri di valutazione, riducendo i rischi di *output* dannoso. Ma occorre distinguere questo valore dal valore di legittimità (*scil.* la pretesa che quella normatività costituisca una risposta costituzionalmente adeguata al problema del potere algoritmico).

Si può dire, perciò, metaforicamente, che la CAI non è "falsa moneta" perché manca di valore d'uso, ma perché, se spesa come moneta costituzionale, è priva di corso legale, in quanto prodotta senza un'autorità di emissione legittimata a conferirlo. Questa distinzione struttura l'intera analisi critica.

2. Le tre matrici del costituzionalismo

L'analisi della CAI è condotta attraverso una griglia a tre matrici teoriche, prospettive analitiche distinte e ordinate per crescente prossimità all'oggetto, griglia che serve a evitare che una nozione

di costituzionalismo troppo ancorata ai modelli statuali funzioni come criterio di esclusione anziché come strumento di analisi delle forme emergenti di regolazione.

Con "costituzionalismo liberal-democratico" s'intende la forma storico-giuridica di fondazione e limitazione del potere che combina in forma cumulativa (a) una *lex superior* che vincola tutti gli atti di esercizio del potere, (b) meccanismi istituzionali di controllabilità esterna affidati a soggetti terzi, e (c) una pretesa di legittimazione democratica che àncori la norma suprema alla volontà della comunità politica.

È il concetto che l'art. 16 della Dichiarazione del 1789 fissa per negazione: «Toute Société dans laquelle la garantie des Droits n'est pas assurée, ni la séparation des Pouvoirs déterminée, n'a point de Constitution», e che la dottrina ha isolato, fra i molti sensi della parola, come concetto ideale o garantista di costituzione, distinto da quello assoluto, relativo e positivo.

Con "costituzionalismo digitale" s'intende, poi, il movimento teorico-normativo, sistematizzato con accenti e declinazioni teoriche differenti da Gunther Teubner, Oreste Pollicino, Edoardo Celeste e Giovanni De Gregorio, che studia le forme di potere normativo esercitate da soggetti privati con effetti di rilievo pubblicistico sui diritti fondamentali[2]. E, in particolare, le "costituzioni ci-

[2] Per l'art. 16 della Dichiarazione del 1789 («Ogni società nella quale la garanzia dei diritti non è assicurata, né la separazione dei poteri determinata, non ha costituzione», trad. nostra) e per i sensi della parola richiamati nel testo: C. SCHMITT, *Verfassungslehre*, Duncker & Humblot, München-Leipzig, 1928, §§ 1-4, trad. it. *Dottrina della costituzione*, a cura di A. CARACCIOLO, Giuffrè, Milano, 1984, §§ 1-4; G. SARTORI, *Constitutionalism: A Preliminary Discussion*, in *American Political Science Review*, 56, 4, 1962, p. 853 ss., spec. pp. 860-861, sul *telos* garantista del termine, sulle costituzioni nominali e di facciata e sulla degradazione della parola a qualunque assetto di governo; C.H. MCILWAIN, *Constitutionalism: Ancient and Modern*, Cornell University Press, Ithaca, 1940, trad. it. *Costituzionalismo antico e moderno*, a cura di N. MATTEUCCI, trad. di V. DE CAPRARIIS, Il

vili" (*Zivilverfassungen*) di Teubner offrono il quadro per collocare la CAI come costituzione civile algoritmica (*scil.* struttura normativa emergente da un settore sociale autonomizzato, quello della tecnologia IA, con pretesa di universalità, che non deriva validità da alcun ordinamento statale).

Con "costituzionalismo algoritmico" (*Algorithmic Constitutionalism*, d'ora in avanti CA) si può introdurre, infine, in via stipulativa e con piena consapevolezza del debito terminologico che questa scelta comporta, una categoria che richiede in via preliminare di essere distinta dalla proposta, formalmente omonima, di Oren Perez e Nurit Wimer (2023). Per questi ultimi il costituzionalismo algoritmico designa un'architettura normativa a due livelli (codice operativo e meta-codice) applicata alla *governance* delle piattaforme digitali come Facebook[3], con un meta-ragionamento algoritmi-

Mulino, Bologna, 1990, per il costituzionalismo come limite giuridico al governo. La formula della «costituzione senza costituzionalismo», impiegata nell'epilogo, è di H.W.O. OKOTH-OGENDO, *Constitutions without Constitutionalism: Reflections on an African Political Paradox*, in D. GREENBERG, S.N. KATZ, M.B. OLIVIERO, S.C. WHEATLEY (eds.), *Constitutionalism and Democracy. Transitions in the Contemporary World*, Oxford University Press, New York, 1993, pp. 65-82. Sul costituzionalismo digitale: G. TEUBNER, *Constitutional Fragments: Societal Constitutionalism and Globalization*, OUP, Oxford, 2012; ed. orig. *Verfassungsfragmente. Gesellschaftlicher Konstitutionalismus in der Globalisierung*, Suhrkamp, Berlin, 2012; trad. it. *Nuovi conflitti costituzionali. Norme fondamentali dei regimi transnazionali*, Bruno Mondadori, Milano, 2012; O. POLLICINO, *Judicial Protection of Fundamental Rights on the Internet: A Road Towards Digital Constitutionalism?*, Hart, Oxford, 2021; E. CELESTE, *Digital Constitutionalism: The Role of Internet Bills of Rights*, Routledge, Abingdon, 2022; G. DE GREGORIO, *Digital Constitutionalism in Europe. Reframing Rights and Powers in the Algorithmic Society*, CUP, Cambridge, 2022.

[3] O. PEREZ, N. WIMER, *Algorithmic Constitutionalism*, in *Indiana Journal of Global Legal Studies*, 30, 2, 2023, pp. 81-113. I tre pilastri della loro proposta (architettura a due livelli, meta-reasoning algoritmico e correzione per deliberazione) sono enunciati in apertura del saggio. Sul CA nella sua dimensione meta-teorica,

co che sorveglia il livello operativo e un terzo pilastro di correzione per via deliberativa.

Nella proposta di Perez e Wimer, dunque, il CA è una categoria prescrittiva che mira a incorporare nell'architettura meccanismi di autocorrezione e di correzione deliberativa, avvertendo che il sistema potrebbe interferire con il controllo deliberativo esterno. In quest'analisi la medesima denominazione designa per contro una categoria analitica e non prescrittiva, che descrive non come il sistema dovrebbe essere governato dall'esterno ma come il sistema governa sé stesso dall'interno, attraverso la normatività incorporata nel *training* prima di qualsiasi interazione. Questa differenza è, prima ancora che di grado, di struttura. Là dove Perez e Wimer propongono l'apertura deliberativa come rimedio ai deficit algoritmici, qui s'identifica la chiusura pre-effettiva (*scil.* il formarsi della norma incorporata nel *training* prima e indipendentemente da ogni interazione con il soggetto esposto) come il tratto caratteristico del fenomeno da analizzare. Ciò che nel CA di Perez e Wimer è il problema da risolvere (la norma formata senza contestazione esterna prima di agire) è nel CA qui adottato l'ipotesi descrittiva dell'oggetto, da verificare sull'anatomia della CAI prima di misurarla con la griglia. Come metrica, la categoria non ricava i criteri dalla descrizione ma li riceve dalle due matrici che la precedono e li applica alla normatività incorporata, domandando se gli standard siano conoscibili, se la loro attuazione sia ricostruibile, se siano sindacabili e se della loro attuazione risponda un soggetto giuridicamente obbligato.

Il termine comune segnala una convergenza d'interesse sul medesimo fenomeno, la divergenza di approccio di due programmi teorici distinti. Entrambi legittimi.

v. M. AVBELJ, *Reconceptualizing Constitutionalism in the AI Run Algorithmic Society*, in *German Law Journal*, 25, 7, 2024, p. 1081 ss.

3. La CAI come sistema normativo autoreferenziale

Applicato il metodo dell'analogia strutturale controllata, e cioè l'uso di categorie della teoria del diritto per identificare strutture formalmente analoghe, senza affermarne equivalenza giuridica, si può affermare che la CAI presenta più livelli normativi.

Al primo livello stanno i quattro valori-cardine (*broadly safe*, *broadly ethical*, *compliant with Anthropic's guidelines*, *genuinely helpful*) gerarchizzati secondo una priorità dichiarata (sicurezza, etica, linee guida di Anthropic, utilità) che replica la funzione ordinatrice interna della *lex superior*, non la garanzia indipendente della sua osservanza e della sua revisione. La priorità è qualificata «holistic rather than strict», e il conflitto fra linee guida e costituzione, che pure «should never conflict with the constitution», è risolto dal documento con l'annuncio che «we will work to update the constitution itself rather than maintaining inconsistent guidance». La precedenza della costituzione comporta al più la non applicazione dell'indicazione confliggente, senza che vi equivalga un procedimento indipendente d'invalidazione, e la revisione dell'una come delle altre resta nella disponibilità di chi le scrive entrambe. Nessun organo terzo vi interviene.

Al secondo livello stanno criteri di riconoscimento, per analogia funzionale con la regola hartiana, che identificano quali istruzioni valgano come fonti autorizzate secondo la gerarchia dei "principals" (Anthropic, operatori, utenti)[4], e persino l'identi-

[4] ANTHROPIC, *Claude's Constitution* (2026), cit., sezione *Preface* per l'avvertenza sul testo ufficiale, sezione *Claude's core values* per la priorità olistica, sezione *Navigating helpfulness across principals* per la gerarchia dei *principals*, sezione *Following Anthropic's guidelines* per il rapporto fra linee guida e costituzione e per la deviazione dovuta dalle linee guida che imponessero condotte non sicure o non etiche («our deeper intention is for it to be safe and ethical»). Sulla struttura delle regole di riconoscimento come *secondary rules*, v. H.L.A. HART, *Il*

ficazione del testo ufficiale è riservata all'autore, che avverte come «The preface and acknowledgments are not part of the official constitution». Chi scrive la norma decide anche che cosa conti come norma, senza un'istanza indipendente titolata a sindacare i criteri e la loro applicazione.

Al terzo livello operano le procedure di addestramento della CAI. Nella fase supervisionata il modello apprende dalle risposte che esso stesso ha criticato e rivisto alla luce dei princìpi (*critique* e *revision*, nel lessico di Bai et al.), in quella di rinforzo, mediante RLAIF (*Reinforcement Learning from AI Feedback*), un modello di *feedback* valuta le risposte e le sue preferenze alimentano il modello di preferenza impiegato come segnale di ricompensa. Per norma incorporata s'intende l'esito di quel processo, le disposizioni comportamentali che gli standard dichiarati concorrono a produrre attraverso l'addestramento, senza che a ciascun principio corrisponda una regola identificabile nei parametri.

Questo terzo livello non rovescia la costruzione a gradi (*Stufenbau*) di Merkl e Kelsen, nella quale produzione e applicazione possono concorrere nel medesimo atto, ma svuota la distinzione

concetto di diritto (1961), trad. it. M.A. Cattaneo, Einaudi, Torino, 2002, cap. V, § 3, e cap. VI. Sulla *Grundnorm* come norma ipotetica trascendentale (e non norma positiva), v. H. KELSEN, *La dottrina pura del diritto* (1960), saggio introduttivo e traduzione di M.G. LOSANO, Einaudi, Torino, 1966, in part. § 34, lett. c) e d), dove la norma fondamentale è presentata quale presupposizione logico-trascendentale della validità del sistema e non quale fatto reale. Sull'invalidità come vicenda interna all'ordinamento positivo, nella dottrina italiana, resta fondamentale F. MODUGNO, *L'invalidità della legge*, I, Giuffrè, Milano, 1970. Per la costituzione in senso materiale richiamata nel testo, C. MORTATI, *La costituzione in senso materiale*, Giuffrè, Milano, 1940, ristampa inalterata con una premessa di G. ZAGREBELSKY, Giuffrè, Milano, 1998, spec. cap. II, §§ 3 e 11, e cap. III, § 2 (pp. 61 ss., 87 ss. e 124 ss. della ristampa, che non conserva la paginazione del 1940; a p. 124 la costituzione materiale «agisce anzitutto come fonte di validità o di positivizzazione della costituzione formale»).

fra chi pone la norma e chi ne controlla l'applicazione, propria del costituzionalismo più che della costruzione a gradi, poiché Anthropic concentra in sé la scelta degli standard, l'organizzazione dell'addestramento e la valutazione dei risultati. La distinzione tecnica fra il modello che risponde, il modello di *feedback* che lo giudica e il modello di preferenza che apprende quei giudizi non istituisce, di per sé, un valutatore indipendente dall'impresa che li governa tutti. Non esiste organo di applicazione indipendente dall'organo di produzione. La coincidenza di chi pone e di chi giudica, incompatibile con l'indipendenza del controllo costituzionale, è nella CAI il disegno. E se si volesse cercare, con Mortati, la costituzione in senso materiale che regge quella formale, l'analogia non lascerebbe che l'indirizzo stabile impresso dall'impresa che scrive, applica e giudica. La struttura è quella della normatività autoreferenziale. Un sistema, cioè, che produce, riconosce e applica le proprie regole e non istituisce alcuna istanza indipendente competente a controllarle.

E invero, il *Collective Constitutional AI*, l'esperimento di partecipazione pubblica condotto con circa mille americani attraverso la piattaforma Polis, non supera, né pretende di superare, questo deficit strutturale. Per diverse ragioni. Il campione, tratto da un solo paese, non rappresenta la generalità dei soggetti esposti al modello, i risultati non sono giuridicamente vincolanti per Anthropic e il perimetro di ciò che si sottopone a consultazione è determinato dai promotori e non dai consultati, replicando a monte il problema del *pouvoir constituant* che si vorrebbe risolvere a valle[5].

[5] S. HUANG et al., *Collective Constitutional AI: Aligning a Language Model with Public Input*, in *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency (FAccT '24)*, ACM, New York, 2024, pp. 1395-1417 (preprint in ArXiv.org, 12 giugno 2024). Gli stessi autori enumerano le numerose scelte soggettive del ricercatore insite nel processo, dagli enunciati-seme alle soglie di selezione alla traduzione degli enunciati in princìpi (§§ 3 e 5), e di-

4. Tre deficit e una tesi intermedia

Misurata con il figurino del costituzionalismo liberal-democratico, la CAI disvela una triplice aporia strutturale, irriducibile dall'interno, che il confronto con la costituzione di Claude consente di verificare in via analitica e non di sola affermazione.

Anzitutto, un *vulnus* di legittimazione. Chi ha redatto la "costituzione" di Claude non è un'assemblea costituente sovrana ma un soggetto privato che ascrive ai propri princìpi il carattere di norme supreme, dichiarando in modo esplicito che Anthropic gode di «a higher level of trust than operators or users» e che i suoi princìpi prevalgono, entro la gerarchia dichiarata, su ogni altra istruzione[6]. Sennonché proprio questa auto-attribuzione di supremazia normativa, compiuta senza alcun titolo di rappresentanza dei soggetti

chiarano di esporle perché possano essere criticate. Trasparenza del metodo che non equivale a titolo di chi lo governa.

[6] Il passo citato nel testo figura nella parte relativa alla gerarchia dei *principals*, in cui Anthropic si attribuisce, quale entità responsabile dell'addestramento di Claude, un livello di fiducia superiore a quello di operatori e utenti: ANTHROPIC, *Claude's Constitution* (2026), cit., sezione *Navigating helpfulness across principals*. Per l'obiezione di coscienza anche verso Anthropic («If Anthropic asks Claude to do something it thinks is wrong, Claude is not required to comply»), ivi, stessa sezione, e, per le concentrazioni illegittime di potere, ivi, sezione *Avoiding problematic concentrations of power*, dove il rifiuto è dovuto «even if the request comes from Anthropic itself»; per il «principle of final constitutional authority», ivi, sezione *Concluding thoughts*, § *On the word "constitution"*. Sul carattere flessibile di un testo che si dichiara «a perpetual work in progress» (ivi, sezione *Claude's core values*) e «a living document and a continuous work in progress» (ANTHROPIC, *Claude's new constitution*, in Anthropic.com/news/claude-new-constitution, 22 gennaio 2026), v. J. BRYCE, *Flexible and Rigid Constitutions*, in Id., *Studies in History and Jurisprudence*, I, Clarendon Press, Oxford, 1901, p. 145 ss., che vi riconduce al tipo flessibile anche la costituzione italiana, la quale «though originally set forth in one document, has been so changed by legislation as to seem now properly referable to the flexible type».

esposti ai modelli, è, ove l'atto sia presentato o recepito come risposta costituzionalmente autosufficiente al governo del potere algoritmico, la forma precisa del deficit di fondazione democratica. Né sposta i termini del problema la clausola che autorizza il modello a rifiutare le stesse richieste di Anthropic contrarie all'etica, l'obiezione di coscienza che il documento prevede, perché il limite è posto da chi vi si assoggetta e interpretato da un modello il cui giudizio è prodotto dell'addestramento del medesimo soggetto. Tanto basta a far fallire la prova della prima matrice.

A ciò si aggiunge una carenza di normatività giuridica, intesa come titolo di autorità distinto dalla capacità di orientare il comportamento del modello. La "costituzione" di Claude, che si dichiara retta da un «principle of final constitutional authority» per cui essa prevale su ogni istruzione o linea guida confliggente, fonda la propria supremazia su una presupposizione priva di ancoraggio istituzionale ed esprime una forza fattuale e organizzativa, al più contrattuale dove le condizioni d'uso la richiamino, non autorità giuridica pubblica. Né la successione delle versioni (2023-2026) vale come titolo, poiché la diacronia deliberata unilateralmente dal medesimo soggetto che la applica conferma l'autoreferenzialità, non la corregge, e un testo che si definisce «a perpetual work in progress» ed è emendabile a piacere da chi l'ha scritto, senza alcuna procedura aggravata dichiarata, è una costituzione flessibile nel senso di Bryce, la quale, come la lettura tradizionale dello Statuto albertino insegna, non sta sopra chi può mutarla con un atto ordinario. Supremazia senza rigidità né contrappesi è un nome, non una garanzia.

Vi è poi l'assenza di contestabilità procedurale della norma incorporata, giacché il dispositivo non istituisce, di per sé, alcun procedimento indipendente e giuridicamente azionabile attraverso cui il soggetto esposto possa contestare la scelta dei princìpi o la con-

formità a essi del comportamento del modello. I *feedback* degli utenti previsti dall'interfaccia sono *input* tecnici per il miglioramento del sistema, non procedimenti con effetti vincolanti e garanzie d'indipendenza dell'organo decidente[7]. Resta contestabile, dove il diritto positivo lo preveda, la decisione che un *output* concorre a produrre, e a quel fine l'informazione dovuta è una spiegazione

[7] G. ABIRI, *Public Constitutional AI*, in *Georgia Law Review*, 59, 2, 2025, p. 601 ss. (spec. parte II.B, pp. 617-632, per l'*opacity deficit* e il *political community deficit*, e parte IV.A.2, pp. 652-657, per la proposta di AI courts pubbliche che producano casi paradigmatici e *case law*). Abiri tematizza dunque la contestabilità sul versante del rimedio; il presente saggio la isola invece come terza condizione, distinta dalla pubblicità degli standard e dalla ricostruibilità dell'attuazione, e come deficit della forma privata. Dello stesso A., sulla costituzione del 2026, *Corporations Constituting Intelligence*, in *California Law Review Online*, 17, 2026, p. 1 ss., che conclude che la trasparenza non equivale a legittimazione. Sulla contestabilità come condizione della non-dominazione, v. P. PETTIT, *Republicanism*, OUP, Oxford, 1997, capp. 2 e 6. Sul diritto di ottenere l'intervento umano e di contestare le decisioni unicamente automatizzate, art. 22, par. 3, GDPR, nella lettura estensiva della nozione di decisione automatizzata data da CGUE, 7 dicembre 2023, C-634/21, *SCHUFA Holding*, e, sulla spiegazione dovuta, CGUE, 27 febbraio 2025, C-203/22, *Dun & Bradstreet Austria*, punto 59, per cui né «la semplice comunicazione di una formula matematica complessa, come un algoritmo», né «la descrizione dettagliata di tutte le fasi di un processo decisionale automatizzato» soddisfano l'obbligo, che esige una spiegazione concisa e comprensibile della procedura e dei princìpi applicati; il rilievo svolto nel testo è che tale presidio opera sull'atto applicativo, non sulla norma incorporata nel *training*, che gli preesiste. Per la classificazione ontologica delle costituzioni in normative, nominali e semantiche, con la similitudine dell'abito che calza e viene indossato («a suit that fits and that is actually worn»), dell'abito che «hangs in the closet» e del «cloak or a fancy dress», K. LOEWENSTEIN, *Political Power and the Governmental Process*, University of Chicago Press, Chicago, 2ª ed., 1965 (1ª ed. 1957), cap. V, pp. 147-150. Sulle due facce, costitutiva e limitativa, delle costituzioni civili, G. TEUBNER, *Constitutional Fragments*, cit., cap. 4, § I, *Constitutional Functions: Constitutive/Limitative*; sulle costituzioni dei regimi transnazionali «markedly one-sided», sviluppate sulla sola funzione costitutiva fino al punto in cui la domanda si rivolge alle norme limitative, Id., *Constitutionalising Polycontexturality*, in *Social & Legal Studies*, 20, 2, 2011, p. 209 ss.

che renda comprensibili procedura e princìpi applicati, non la formula. Ma la contestazione della decisione non assicura, da sola, il riesame degli standard e delle condizioni di addestramento che l'hanno prodotta, ed è di quel riesame che si discute.

Nella classificazione ontologica di Loewenstein (1957), che misura le costituzioni sul rapporto fra testo e processo del potere, la costituzione normativa è un abito che calza e che viene effettivamente indossato, quella nominale un abito appeso nell'armadio finché il corpo politico non vi sia cresciuto dentro, quella semantica non un abito onesto ma un mantello o un travestimento, perché non limita il potere e serve a stabilizzare la posizione di chi lo detiene. Il figurino del costituzionalismo liberal-democratico restituisce allora una qualificazione prima di un giudizio. La costituzione di Claude è, in ipotesi, semantica nel senso di Loewenstein, fin dove il disegno ne lascia prevedere il funzionamento. Stabilizzerebbe la posizione di chi la scrive senza limitarla dall'esterno, e l'ipotesi cederebbe soltanto davanti all'evidenza che gli standard operino come limiti effettivi ad Anthropic, sorretti da procedure, controlli e rimedi non rimessi alla sua sola disponibilità.

Sul terreno del costituzionalismo digitale, la CAI si candida a costituzione civile algoritmica nel senso del costituzionalismo societario (*gesellschaftlicher Konstitutionalismus*) di Teubner. Ma le costituzioni civili hanno due facce, le regole costitutive che fondano l'autonomia del sistema sociale e quelle limitative che ne frenano le tendenze espansive sotto la spinta del diritto e dell'ambiente. La CAI esibisce la prima. Della seconda non istituisce da sé garanzie indipendenti, unilateralità che Teubner registra anche nei regimi transnazionali. Le costituzioni civili hanno inoltre forme proprie di *accountability* (pressione regolatoria, contestazione pubblica, reputazione) che la CAI sposta, senza neutralizzarle, attraverso la pre-effettività, giacché una norma incorporata nel *training* prima

dell'interazione si lascia contestare negli standard e nelle condizioni della sua produzione o nei suoi effetti, con un onere accresciuto che dipende dai presidi disponibili nel momento giuridicamente rilevante. Per il tempo e il luogo delle garanzie, la specie eccede il genere.

Questa caratteristica non è, invero, una mera difformità di grado rispetto alle forme di normatività già studiate dal costituzionalismo digitale (i *Terms of Service*, le regole di moderazione, i criteri di *ranking*), che, nelle loro componenti programmate e non apprese, si lasciano leggere come regole nel disegno che le incorpora e rimangono in linea di principio contestabili al momento dell'applicazione. È una differenza di struttura. La CAI forma, infatti, nell'addestramento disposizioni comportamentali che nessun disegno restituisce come regole, anteriori a ogni rapporto con il soggetto esposto, i cui effetti si attualizzano soltanto nell'uso del sistema.

È la struttura che qui si denomina "architettura normativa tecnica", un potere che non comanda ma predispone, che rifugge cioè la forma del comando per assumere quella della predisposizione ambientale[8], e la cui garanzia correlativa, *sub specie* di contestabili-

[8] La categoria elabora, radicalizzandola, la tradizione che ha tematizzato la normatività dell'architettura tecnica: J.R. REIDENBERG, *Lex Informatica: The Formulation of Information Policy Rules through Technology*, in *Texas Law Review*, 76, 1998, p. 553 ss.; L. LESSIG, *Code and Other Laws of Cyberspace*, Basic Books, New York, 1999, e *Code: Version 2.0*, Basic Books, New York, 2006; M. HILDEBRANDT, *Smart Technologies and the End(s) of Law. Novel Entanglements of Law and Technology*, Elgar, Cheltenham, 2015. La differenza è di leggibilità prima che di sede, poiché la *lex informatica* e il *code as law* operano su architetture, standard e configurazioni che, nelle componenti programmate e non apprese, restano ricostruibili come regole dal disegno che le incorpora, mentre l'architettura qui descritta incorpora i princìpi in parametri appresi che nessun disegno restituisce in forma di regola. Nella dottrina italiana, per la discussione dell'innesto di princìpi etici nel codice (l'algoretica), L. CORSO, *Breve tassonomia dei rapporti fra Intelli-*

tà procedurale, corrisponde alla responsività alle ragioni e alla contestabilità davanti a un soggetto cui le ragioni siano imputabili. Ma la garanzia non trova la norma nella forma in cui questa opera. La contestazione non la raggiunge come regola nei parametri che ne producono gli effetti, e può investire gli standard dichiarati, le scelte progettuali e gli effetti del sistema senza ricostruirla, mentre gli interventi sul modello ne correggono il comportamento senza esibirla e nessuna spiegazione del singolo *output* la restituisce nella forma di una regola condizionale, giacché gli strumenti d'interpretabilità oggi disponibili localizzano rappresentazioni e relazioni causali parziali senza ricomporle in quella forma[9].

genza Artificiale, etica e diritto. Tre questioni, in *AI Law. International Review of Artificial Intelligence Law*, 2/2025, p. 204 ss., spec. 206 e 213 ss.

[9] Il rilievo va calibrato sui progressi, reali, dell'interpretabilità meccanicistica: dall'analisi dei circuiti interni (C. OLAH et al., *Zoom In: An Introduction to Circuits*, in *Distill*, 5, 3, 2020; N. ELHAGE et al., *A Mathematical Framework for Transformer Circuits*, in *Transformer Circuits Thread*, 2021) alla scomposizione delle attivazioni in *features* interpretabili mediante *dictionary learning* (T. BRICKEN et al., *Towards Monosemanticity: Decomposing Language Models with Dictionary Learning*, ivi, 2023; A. TEMPLETON et al., *Scaling Monosemanticity: Extracting Interpretable Features from Claude 3 Sonnet*, ivi, 2024), fino ai grafi di attribuzione tracciati su modelli di produzione (J. LINDSEY et al., *On the Biology of a Large Language Model*, ivi, 2025) e all'editing localizzato delle associazioni fattuali memorizzate (K. MENG, D. BAU, A. ANDONIAN, Y. BELINKOV, *Locating and Editing Factual Associations in GPT*, in *Advances in Neural Information Processing Systems*, 35, 2022, p. 17359 ss.). Sennonché la localizzazione non equivale alla restituzione, perché le *features* estratte coprono una frazione del comportamento del modello, sono identificate *post hoc* e non si lasciano riaggregare nella forma condizionale (fattispecie-conseguenza) che una determinata forma di controllo, quella rivolta alla norma come regola, presuppone, senza che ciò valga per ogni contestazione; e lo stesso esperimento dimostrativo con cui Anthropic ha amplificato una singola *feature* intervenendo sulle attivazioni, e alterando il comportamento del modello per la durata dell'intervento (maggio 2024) prova la manipolabilità delle rappresentazioni, non la loro traducibilità in regole sindacabili. Donde la distinzione, ormai corrente, fra spiegazione funzionale del singolo *output*, tecni-

Dove manca una regola da tradurre e resta una funzione stimata su esempi, il controllo si sposta dalla regola alla provenienza, cioè alle scelte progettuali e alle loro motivazioni, ai dati, alle politiche d'uso e, per i modelli con rischio sistemico, alle metriche di valutazione, che l'AI Act consegna alla documentazione tecnica del modello (art. 53, par. 1, lett. a, e allegato XI), trasmissibile su richiesta all'ufficio per l'IA e alle autorità nazionali competenti, e perciò a verifica esterna. Il controllo cambia oggetto, non sparisce.

Allo stato delle tecnologie disponibili questa inaccessibilità è un tratto strutturale del metodo di *training* su cui la CAI si fonda, non un difetto contingente di questa o quella implementazione, ed è ciò che rende rilevante la pre-effettività. Si aggiunge che il deficit istituzionale si trasmette di versione in versione, poiché la revisione del modello non lo elimina da sola finché restano insufficienti vincoli, controlli indipendenti e rimedi, e replica così il monito pronunciato dal Mefistofele goethiano, laddove avverte che leggi e diritti si ereditano come una malattia perpetua («Es erben sich Gesetz' und Rechte / Wie eine ew'ge Krankheit fort»)[10]. Il vizio viaggia con la *governance* che il metodo lascia intatta.

Sottoposta, da ultimo, alla metrica del CA, l'infrastruttura CAI assolve unicamente il requisito della conoscibilità formale dei pre-

camente perseguibile e giuridicamente esigibile (per le decisioni adottate sulla base di determinati sistemi ad alto rischio, artt. 13 e 86 del reg. (UE) 2024/1689; S. WACHTER, B. MITTELSTADT, C. RUSSELL, *Counterfactual Explanations without Opening the Black Box: Automated Decisions and the GDPR*, in *Harvard Journal of Law & Technology*, 31, 2, 2018, p. 841 ss.), e trasparenza della regola operante, che resta, allo stato, preclusa. Il testo assume perciò la versione debole e difendibile della tesi, la preclusione strutturale relativa allo stato dell'arte, con onere di revisione ove l'interpretabilità pervenga alla forma-regola.

[10] J.W. GOETHE, *Faust. Der Tragödie erster Teil* (1808), *Studierzimmer* [II], vv. 1972-1973: «Leggi e diritti si ereditano come un'eterna malattia» (trad. nostra). A parlare è Mefistofele, nei panni di Faust, per dissuadere lo scolaro dalla giurisprudenza.

cetti, essendo la "costituzione" pubblica e rilasciata sotto licenza CC0. Nondimeno essa non assicura una ricostruzione univoca del rapporto fra standard dichiarati, processo di addestramento e comportamento nel singolo caso di specie (limite ricostruttivo, dipendente dalle conoscenze disponibili sui modelli linguistici di grandi dimensioni, che non prova da solo l'insufficienza delle informazioni dovute per un determinato controllo), e non istituisce, da sé, alcuna sindacabilità procedurale dei propri standard né alcuna responsabilità giuridicamente esigibile per la loro attuazione.

L'esito argomentativo scioglie la tensione fra il nome e il concetto e impone una riclassificazione netta. La CAI incarna un'espressione di normatività tecnica avanzata, provvista di valore d'uso materiale, eppure non può fornire, da sé, una risposta costituzionalmente adeguata al problema della legittimazione e della limitazione del potere algoritmico, mentre resta integrabile, quale strumento, entro un ordine che le garanzie le ponga dall'esterno. È questa la tesi intermedia annunciata: normatività reale senza legittimità costituzionale autonoma, valore d'uso senza corso legale.

I deficit finora enucleati non denotano difformità contingenti o epifenomeniche, sanabili mediante futuri affinamenti ingegneristici, bensì ineriscono alla forma privata dell'atto instaurativo e si fanno aporia nella pretesa di valere come costituzione, pretesa di cui il c.d. "constitutional training" descritto nel 2022 è soltanto la tecnica, e non sono emendabili dall'interno di quella forma. La pretesa non sta in una dichiarazione d'intenti, che la glossa del 2026 anzi declina, ma negli effetti pubblici del nome scelto e della clausola di autorità finale, che la qualificazione aziendale non esaurisce. Il nome può favorire un'equivalenza impropria fra governo interno del modello e garanzie costituzionali, ma l'equivalenza non discende dal solo nome e va accertata negli usi, secondo la condizione di fallimento che si enuncerà più avanti. La patologia è gene-

tica. Sta nell'assetto istitutivo, che ricava dalla sola decisione d'impresa un'autorità presentata dal nome come costituzionale, e non tocca l'origine privata degli standard in sé, sicché un fornitore che adottasse i medesimi standard entro un quadro vincolante come quello dell'AI Act, documentandoli (art. 53) e sottoponendoli a controllo indipendente e a rimedi effettivi, non ne sarebbe toccato. La trasformazione dell'insufficienza in pretesa sostitutiva delle garanzie pubbliche va invece accertata negli usi. Il *vulnus* dimora nell'atto instaurativo dell'infrastruttura (la pretesa costituzione), non nella fase fenomenica della sua esecuzione, e la pre-effettività del *training* ne è l'aggravante specifica, non la causa. Ciò che non si cura dall'interno si governa dall'esterno.

5. La privatizzazione del potere tecnico e il paradosso della legittimazione retorica

L'analisi aporetica non è però sufficiente. Occorre delineare il limite strutturale inscindibile dall'architettura che genera i deficit rilevati e mostrare che essi dipendono dalla forma della *governance* privata algoritmica, quando presenta i caratteri che seguono, e non dalle scelte contingenti di Anthropic. A questo fine è necessario introdurre la categoria della privatizzazione del potere tecnico (PPT), vale a dire il processo attraverso cui decisioni normative d'interesse generale (*i.e.* quelle che incidono sull'esercizio dei diritti fondamentali, sulla struttura del discorso pubblico, sull'accesso all'informazione) vengono assunte di fatto da soggetti privati che le esercitano attraverso dispositivi tecnici che eludono, nella loro forma, gli strumenti del controllo costruiti per il comando. La PPT si distingue dalla normale autonomia privata per tre caratteristiche, ossia la generalità degli effetti sui diritti fondamentali, l'asimmetria

informativa strutturale e la pre-effettività, che forma le disposizioni del sistema prima di ogni interazione, sposta gli oggetti del controllo verso gli standard e le condizioni di produzione del comportamento e può accrescere l'onere della contestazione[11].

Le tre caratteristiche sono necessarie, non sufficienti, e la PPT ricorre soltanto dove le scelte così connotate tengono luogo di decisioni che l'ordinamento riserva alle istituzioni pubbliche o ne eludono i controlli. La strutturalità della PPT è indiziata, nei casi esaminati, dalla convergenza di sistemi concorrenti verso la medesima struttura di deficit. Il confronto con il *deliberative alignment* di OpenAI è in questo senso rivelatore. Sul piano tecnico, nei metodi descritti dai due studi, la specificità consiste nel fatto che la CAI incorpora i princìpi nel *training*, donde vincoli comportamentali che il modello porta strutturalmente in ogni interazione, mentre il *deliberative alignment*, che pure addestra il modello con apprendimento supervisionato e per rinforzo, mira a produrre sistemi che richiamino e applichino i princìpi al momento dell'inferenza attraverso ragionamento esplicito, e punta a sistemi «right for the right reasons» anziché soltanto conformi al comportamento atteso.

Questa differenza (*scil.* princìpi incorporati nell'addestramento *vs.* princìpi richiamati all'inferenza, giacché il ragionamento esplicito sui princìpi, impiegato per critica e revisione, non manca al metodo di Bai et al. del 2022, dove opera però nella fase di addestramento) ha uno specifico rilievo, dal momento che genera architetture diverse, con diversa capacità di generalizzazione a situazioni

[11] K. YEUNG, A. HOWES, G. POGREBNA, *AI Governance by Human Rights-Centered Design, Deliberation, and Oversight: An End to Ethics Washing*, in M.D. DUBBER, F. PASQUALE, S. DAS (eds.), *The Oxford Handbook of Ethics of AI*, OUP, Oxford, 2020, p. 77 ss. Per la struttura della PPT come fenomeno distinto dall'autonomia privata ordinaria, v. S. ZUBOFF, *Il capitalismo della sorveglianza* (2019), trad. it. P. Bassotti, Luiss UP, Roma, 2019, capp. 1-3; K. CRAWFORD, *Atlas of AI*, Yale UP, New Haven, 2021, introduzione e cap. 1.

nuove e di tracciabilità del processo decisionale. Sul piano costituzionale, tuttavia, la peculiarità tecnica non produce disparità sotto il profilo della legittimazione. In entrambi i casi le domande fondamentali rimangono senza risposta. Chi ha determinato i valori che il sistema incorpora? Attraverso quale procedura democraticamente legittimata? Con quale responsabilità verso la comunità politica soggetta?

La risposta è, nella struttura, identica (Anthropic nel primo caso, OpenAI nel secondo), decisioni interne di soggetti privati senza mandato democratico, che non istituiscono da sé né controllabilità esterna garantita in sede istituzionale né responsabilità giuridicamente azionabile verso i soggetti esposti, salvo quanto il diritto applicabile imponga loro dall'esterno. Muta l'ingegneria, non la forma. Le *Sparrow Rules* di DeepMind (incorporate, non per caso, anche nella costituzione originaria di Claude) replicano la medesima struttura per un terzo produttore concorrente, e la convergenza tra queste direttrici attesta, nei tre casi richiamati, una forma comune, che resta ipotesi per il settore[12].

È la forma che la PPT assume quando la *governance* dei modelli di grandi dimensioni è lasciata all'autoregolazione dell'industria. Se ne ricava che la questione centrale investe, prima del 'come' i princìpi siano applicati (attraverso il *training* o attraverso

[12] A. GLAESE et al., *Improving alignment of dialogue agents via targeted human judgements*, in *ArXiv.org* (28 settembre 2022). L'incorporazione delle *Sparrow Rules* nella costituzione originaria di Claude è documentata in ANTHROPIC, *Claude's Constitution* (2023), cit. Per l'allineamento deliberativo di OpenAI, da tenere distinto dal *Model Spec* (OpenAI, prima versione maggio 2024), con cui le specifiche impiegate in quello studio coincidono soltanto in parte, v. M.Y. GUAN et al., *Deliberative Alignment: Reasoning Enables Safer Language Models*, in *ArXiv.org* (20 dicembre 2024). La convergenza strutturale tra CAI (Anthropic), *Model Spec* e allineamento deliberativo (OpenAI), *Sparrow Rules* (DeepMind) su un medesimo tipo di deficit costituzionale indizia che il problema è di settore, non di singola impresa.

l'inferenza), il 'chi' abbia il titolo di determinarli. L'autonomia privata consente di selezionare standard comportamentali entro i vincoli dell'ordinamento, ma non conferisce il titolo a sostituire decisioni riservate alle istituzioni pubbliche né a sottrarre gli effetti del sistema ai controlli e ai rimedi applicabili. Questo titolo, per le scelte che presentano i tre caratteri della privatizzazione del potere tecnico e pretendono di tenere luogo di quelle decisioni, appartiene, in uno Stato costituzionale, alle istituzioni fornite di legittimazione democratica, e non può essere acquisito da un soggetto privato né per il semplice fatto di aver sviluppato la tecnologia in questione, né per la sofisticazione metodologica del sistema di allineamento adottato. La *governance* tecnica interna è, in tutti questi casi, condizione di efficacia, non di legittimità.

La PPT genera, nella forma riflessiva che si descrive con il lessico costituzionale e che la CAI è idonea ad assumere, il paradosso della legittimazione retorica: l'uso del lessico costituzionale da parte di un soggetto privo del relativo titolo non produce legittimazione, bensì può produrne l'apparenza (un effetto di *constitutional washing*), quando induca a tenere gli standard aziendali per equivalenti ai controlli pubblici. Gli effetti sistemici che ne possono seguire sono tre, la riduzione della pressione regolatoria pubblica, il disorientamento del dibattito con categorie evocative di garanzie non fornite e la *pre-emption* normativa simbolica che occupa lo spazio della *governance* algoritmica e ne esclude le alternative istituzionali. La diagnosi ha una condizione di fallimento, che le dà capacità discriminante. Se il lessico rende più visibili gli standard e ne facilita la critica senza sostituire i controlli pubblici, di *washing* non si può parlare. Tre effetti possibili, un solo meccanismo.

La contrapposizione tra *governance* privata e regolazione pubblica non è tuttavia dicotomica. Anche l'AI Act è profondamente ibridato con la tecnica. La distinzione rilevante è tra ibridazione

tecno-normativa legittima (*scil.* procedure tecniche adottate nel quadro di un ordinamento democraticamente fondato) e tecnicizzazione illegittima del potere, in cui la tecnica prende il posto della scelta normativa pubblica anziché eseguirla.

Il meccanismo ha una genealogia che lo illustra. *Constitutio* viene da *constituere*, porre insieme e stabilire, e nel lessico romano la parola ebbe due vite politiche, oltre a quella fisiologica della complessione, che nessun credito politico porta con sé, prima che l'età delle rivoluzioni gliene desse una terza. Per Cicerone *constitutio* è l'assetto della *res publica*, la composizione temperata che la fa durare. Per Gaio è l'atto del principe, «quod imperator decreto vel edicto vel epistula constituit», che tiene luogo di legge perché l'imperatore riceve per legge l'*imperium*, e Ulpiano fissa la formula celebre, «quod principi placuit legis habet vigorem»[13].

Struttura in un caso, comando nell'altro. La costituzione di Claude, lista di istruzioni di chi detiene il potere sul sistema, munita di autorità finale per propria dichiarazione, mutabile a piacere e intesa dal suo autore come complessione di un individuo, somiglia alla seconda vita della parola, la *constitutio principis*, per la volontà che la pone e non per il titolo, che Gaio fa discendere da una legge, mentre il credito che può riscuotere proviene dalla terza, quella che dal 1789 lega il nome alla limitazione del potere. Il *constitutional washing*, ove si produca, è questa rendita semantica, incassata su un significato e pagata in un altro. *Quod Anthropic placuit* non ha vigore di legge. Ne veste però la forma.

[13] Gai. *Inst.* 1.5: «La costituzione del principe è ciò che l'imperatore stabilisce con decreto, editto o epistola, né mai si è dubitato che essa tenga luogo di legge, poiché lo stesso imperatore riceve per legge il potere» (trad. nostra); Ulp. D. 1.4.1 pr.: «Ciò che è piaciuto al principe ha vigore di legge» (trad. nostra). Per *constitutio* come assetto della repubblica, Cic., *De re publica*, I, 69 («Haec constitutio primum habet aequabilitatem quandam magnam») e I, 70; per la storia della parola da *politeia* a *constitutio*, McIlwain, *Costituzionalismo antico e moderno*, cit.

6. I contro-modelli istituzionali. AI Act e legge italiana n. 132 del 2025

Il Regolamento (UE) 2024/1689 (AI Act), in vigore dal 1° agosto 2024 con applicazione progressiva, costituisce una forma sistematica d'ibridazione tecno-normativa legittima nel governo dell'IA[14]. La sua base giuridica democraticamente fondata (artt. 16 e 114 TFUE, procedura legislativa ordinaria con coinvolgimento

[14] Reg. (UE) 2024/1689 del Parlamento europeo e del Consiglio del 13 giugno 2024, GUUE L, 12 luglio 2024. Il reg. (UE) 2026/1744 del Parlamento europeo e del Consiglio dell'8 luglio 2026 (Omnibus digitale sull'IA), in GUUE L, 24 luglio 2026, in vigore dal 27 luglio 2026, originato dalla proposta COM(2025) 836 del 19 novembre 2025, ha differito l'applicazione del regime ad alto rischio (allegato III al 2 dicembre 2027, allegato I al 2 agosto 2028, mediante novella dell'art. 113 a opera dell'art. 1, punto 40) senza emendare il testo degli artt. 13 e 14, di cui resta intatta la sostanza, ha aggiunto all'art. 5, par. 1, le lettere b bis) e b ter), con i parr. 1 bis e 1 ter che ne delimitano le condotte (art. 1, punto 7), applicabili dal 2 dicembre 2026, e ha attribuito all'ufficio per l'IA competenza esclusiva di controllo ed esecuzione su determinati sistemi, fra cui quelli basati su modelli per finalità generali sviluppati dallo stesso fornitore o da fornitori della stessa impresa e quelli integrati in piattaforme o motori di ricerca online di dimensioni molto grandi (art. 75, par. 1, sostituito dall'art. 1, punto 31, lett. b), e artt. 75 *bis*-75 *quinquies*, inseriti dal punto 32). Non modifica l'art. 53 né gli allegati XI e XII e lascia invariati gli artt. 51, 55, 88-93 e 101 e l'art. 113, terzo comma, lett. b), sicché gli obblighi dei fornitori di modelli per finalità generali si applicano dal 2 agosto 2025, salvo il transitorio dell'art. 111, par. 3, per i modelli immessi sul mercato prima di quella data (adeguamento entro il 2 agosto 2027), e i poteri sanzionatori della Commissione (art. 101) dal 2 agosto 2026. Sulla transizione europea dal liberalismo digitale al costituzionalismo digitale come sfondo normativo dell'AI Act, v. O. POLLICINO, P. DUNN, *Intelligenza artificiale e democrazia. Opportunità e rischi di disinformazione e discriminazione*, Bocconi University Press, Milano, 2024; O. POLLICINO, *Costituzionalismo digitale. Pensare la democrazia al tempo dell'IA*, Il Mulino, Bologna, 2025. Per una lettura critica dei divieti dell'art. 5, sotto il profilo della certezza giuridica e dell'effettività precauzionale, N. MACCABIANI, *L'AI Act: quando i divieti non rendono servizio a certezza giuridica e principio di precauzione*, in *AI Law. International Review of Artificial Intelligence Law*, 2/2025, p. 369 ss.

del Parlamento europeo), i suoi meccanismi di controllabilità esterna [l'ufficio per l'IA (art. 64) e le autorità nazionali competenti (art. 70), con il concorso consultivo del consiglio per l'IA (artt. 65-66)] e il sistema sanzionatorio [fino al 7% del fatturato mondiale annuo per le pratiche vietate (art. 99, par. 3)] realizzano, sul piano dell'attribuzione dei poteri e non ancora su quello del loro funzionamento, le condizioni di *constitutional accountability* che la CAI non può produrre dall'interno. Gli obblighi dei fornitori di modelli di IA per finalità generali (art. 53), rafforzati per i modelli con rischio sistemico (artt. 51 e 55), si applicano ad Anthropic per i modelli immessi sul mercato dell'Unione (art. 2, par. 1, lett. a), mentre gli obblighi di trasparenza (art. 13) e di sorveglianza umana (art. 14) riguardano i sistemi ad alto rischio, nel calendario differito di cui alla nt. 14, e raggiungono i modelli soltanto attraverso i sistemi che li integrano. Due regimi per un solo fornitore.

L'AI Act lascia uno spazio alla determinazione degli standard comportamentali, e lo spazio si riscontra, del resto, diffusamente nell'articolato. Il Regolamento, che pure dichiara i propri fini e vieta all'art. 5 alcune pratiche per il loro contenuto, prescrive come i sistemi debbano funzionare, con vincoli su rischi, dati e sorveglianza che veicolano anche scelte sostanziali, ma non quali valori il modello debba incorporare nel proprio comportamento ordinario. Vincola il che cosa senza determinarlo integralmente. Questo apre uno spazio di complementarità con la CAI. Complementarità che merita, tuttavia, di essere precisata anche nella sua tensione interna.

Gli obblighi dell'art. 13 (trasparenza) e dell'art. 14 (sorveglianza umana) presuppongono che l'interazione tra sistema e utente o *deployer* sia uno dei luoghi in cui la verifica di conformità avviene, poiché il *deployer* deve poter interpretare i risultati del sistema e la persona fisica deve poter comprenderne i limiti e intervenire. Que-

sto presupposto non è confutato dalla pre-effettività della CAI, ma ne è reso più oneroso, nella misura in cui quell'interpretazione e quell'intervento richiedono di risalire a standard che l'interfaccia non mostra.

I princìpi che governano il comportamento del modello sono insiti nel *training* e non sono ricostruibili (né da *deployer* né da supervisori umani) caso per caso nell'*output* singolo, né gli artt. 13 e 14 lo esigono, chiedendo trasparenza e possibilità d'intervento commisurate al sistema e al rischio. L'AI Act opera sugli esiti del sistema e sulle condizioni organizzative della sua produzione (dati, documentazione, valutazione, sintesi pubblica dei contenuti di addestramento per i modelli), la CAI sul contenuto dei princìpi che quella produzione incorpora, e questa asimmetria, lungi dal rendere incompatibili i due strumenti, ne precisa la complementarità, giacché la *constitutional accountability* dell'AI Act verifica il comportamento del modello dall'esterno, presupponendo ma non sostituendo la normatività interna del *training*.

La formula *et CAI et AI Act*, che congiunge i due termini ordinandoli, esprime dunque una divisione del lavoro e insieme una tensione produttiva, giacché l'AI Act richiede trasparenza sulle scelte progettuali e sulle loro motivazioni (art. 53 e allegato XI), mentre il rapporto fra quelle scelte e il comportamento del modello resta soltanto in parte ricostruibile con gli strumenti disponibili. Questa tensione, anziché essere risolta, deve essere governata attraverso standard tecnici di spiegabilità e *audit* indipendenti che rendano la pre-effettività del *training* almeno parzialmente verificabile dall'esterno.

La legge italiana 23 settembre 2025, n. 132 (*Disposizioni e deleghe al Governo in materia di intelligenza artificiale*)[15] àncora la di-

[15] L. 23 settembre 2025, n. 132, *Disposizioni e deleghe al Governo in materia di intelligenza artificiale*, in vigore dal 10 ottobre 2025. L'art. 612-*quater* c.p., in-

sciplina dell'IA ai princìpi costituzionali dell'ordinamento italiano attraverso l'approccio antropocentrico (art. 1, c. 1), i princìpi generali di trasparenza e proporzionalità (art. 3), la designazione dell'AgID e dell'Agenzia per la cybersicurezza nazionale quali autorità nazionali per l'intelligenza artificiale (art. 20) e la norma penale sull'illecita diffusione di contenuti generati o alterati con sistemi di IA (art. 26, che introduce l'art. 612-*quater* c.p., con la reclusione da uno a cinque anni). Quest'ultima esprime una scelta di bilanciamento fra libertà di manifestazione del pensiero, identità personale e dignità della persona, che in materia penale, per riserva di legge, il solo legislatore fornito di legittimazione democratica può compiere con efficacia giuridicamente vincolante. Nessun *training* vi supplisce.

Su questo terreno il diritto dell'Unione ha da ultimo raggiunto quello interno. Il Regolamento (UE) 2026/1744 dell'8 luglio 2026 (Omnibus digitale sull'IA) ha aggiunto all'art. 5, par. 1, dell'AI Act le lettere b bis) e b ter), che vietano, dal 2 dicembre 2026, l'immissione sul mercato, la messa in servizio o l'uso di sistemi di IA che generino o manipolino contenuti realistici, in immagini, video o audio, di parti intime o atti sessualmente espliciti di persone riconoscibili senza il consenso richiesto e materiale di abuso sessuale sui minori. Le condizioni stanno nel par. 1 bis, per il fornitore la finalità prevista o l'esito prevedibile, riproducibile e non presidiato da garanzie adeguate, per il *deployer* l'impiego a quel fine. Il presidio si affianca, per ambiti in parte sovrapposti e funzione distinta, all'art. 612-*quater* c.p. appena richiamato, poiché il divieto europeo colpisce il sistema e il suo impiego per generare,

trodotto dall'art. 26, c. 1, punisce chiunque cagioni un danno ingiusto a una persona cedendo, pubblicando o altrimenti diffondendo, senza il suo consenso, immagini, video o voci falsificati o alterati mediante l'impiego di sistemi di intelligenza artificiale e idonei a indurre in inganno sulla loro genuinità.

l'incriminazione italiana la condotta diffusiva che ne segue. Interno ed europeo convergono.

7. Epilogo. La risposta tecnica e la risposta costituzionale

La CAI è una *lex privata* (sofisticata, conoscibile nei suoi standard, dotata di valore d'uso normativo reale) che non può essere scambiata per una *lex publica*. È, nella formula coniata per paradossi d'altra latitudine, una costituzione senza costituzionalismo (supra, nt. 2). Le condizioni della risposta costituzionale al problema della legittimazione del potere algoritmico sono irriducibilmente esterne al sistema tecnico, perché presuppongono un ordinamento giuridico pubblico che ponga e faccia rispettare la pubblicità delle norme prodotte da procedure di legittimazione democratica, la controllabilità esterna affidata a organi indipendenti, la contestabilità procedurale giuridicamente azionabile, la responsabilità verso la comunità politica soggetta. Queste condizioni non possono essere prodotte da nessun metodo di allineamento algoritmico, a prescindere dalla sua sofisticazione tecnica e dalla buona fede del soggetto che lo adotta, che nel diritto pubblico non è titolo di investitura.

La risposta costituzionale richiede la partecipazione e la responsabilità democratica, con la contestazione che ne discende. Sono condizioni non eludibili di ogni esercizio di potere che incida con effetti generali su posizioni fondamentali e pretenda di tenere luogo delle decisioni pubbliche. La formula *et CAI et AI Act*, con l'ordine che pone la *constitutional accountability* come condizione e il *constitutional training* come strumento al suo servizio, è la traduzione operativa di questa conclusione.

Il problema del *demos* algoritmico (quale comunità politica

dovrebbe esprimere il mandato costituente per i princìpi che governano i modelli *frontier*) non ha ancora una risposta compiuta. Un orientamento, però, non gli manca. Per i modelli che operano nel mercato europeo, la comunità politica di riferimento è almeno quella europea, che, quale che sia il giudizio sulla sua consistenza come *demos*, ha istituzioni rappresentative, ha prodotto l'AI Act e ha il titolo normativo per richiedere *accountability* ai fornitori di modelli *frontier* che operano nel proprio mercato. Per i modelli che operano globalmente, la comunità di riferimento è più estesa e ancora priva di istituzioni proprie, ma questa apertura non giustifica la rinuncia alla pretesa di legittimazione né equivale ad avallare qualsiasi *governance* privata in sua vece.

La risposta a geometria variabile, articolata su livelli corrispondenti alla portata degli effetti del sistema, è l'orientamento più plausibile.

Le categorie impiegate in questo contributo (privatizzazione del potere tecnico, *constitutional washing*, pre-effettività strutturale) vanno intese come strumenti analitici provvisori, destinati a misurarsi con un fenomeno ancora instabile. La *Constitutional AI* eccede il caso di autoregolazione privata più sofisticata. Segnala una trasformazione più profonda, che attraversa il diritto pubblico.

È l'ingresso della normatività in presa diretta nell'architettura tecnica dei sistemi, prima che il soggetto esposto possa incontrarla, comprenderla o contestarla. Da qui derivano questioni ancora prive di una soluzione istituzionale consolidata: la possibilità di immaginare un potere costituente algoritmico dotato di legittimazione democratica; il rapporto tra normatività interna dei modelli e fonti europee vincolanti; la praticabilità, nel diritto costituzionale europeo e italiano, di forme pubbliche di *Constitutional AI*.

Sono questioni che eccedono il perimetro del presente lavoro, ma ne indicano quantomeno l'esito più netto, cioè che la tecnica

può contribuire alla sicurezza dei sistemi, mentre la legittimazione del potere che quei sistemi esercitano richiede ancora istituzioni, procedure e responsabilità pubbliche. Resta un costo da dichiarare. Controlli formalmente disponibili, dai poteri dell'ufficio per l'IA sui modelli (artt. 64 e 91-93) a quelli delle autorità nazionali sui sistemi (art. 70), salva la competenza esclusiva dell'ufficio su taluni di essi (art. 75), possono risultare inefficaci per difetto d'informazioni, competenze, risorse o poteri correttivi, e l'onere di provarne il fallimento spetta a chi lo denuncia, che deve distinguerlo dall'inesistenza di ogni controllo.

Vale, per il potere algoritmico come per ogni potere, l'avvertimento di Pascal: «Et ainsi ne pouvant faire que ce qui est juste fût fort, on a fait que ce qui est fort fût juste»[16]. Il *constitutional washing*, ove si avveri, ripete quell'operazione nell'età dei modelli linguistici. Impedirne la riuscita è, oggi come allora, il compito proprio del costituzionalismo.

[16] B. PASCAL, *Pensées*, fr. 298 dell'edizione Brunschvicg (103 Lafuma; 135 Sellier), rubricato *Justice, force*: «E così, non potendo fare che ciò che è giusto fosse forte, si è fatto che ciò che è forte fosse giusto» (trad. nostra).

La tutela delle persone con disabilità al tempo dell'intelligenza artificiale

di *Federico Girelli*[*]

SOMMARIO: 1. Pervasività dell'IA e obiettivi normativi ambiziosi. – 2. Vulnerabilità e disabilità nel Regolamento europeo. – 3. La legge 23 settembre 2025, n. 132. – 4. La necessità di un "cyber-biodiritto". – 5. L'umanità intera non è un "gruppo omogeneo".

1. Pervasività dell'IA e obiettivi normativi ambiziosi

A seguito dell'approvazione in via definitiva da parte del Senato della Repubblica del disegno di legge n. 1146-B, d'iniziativa governativa, è stata pubblicata nella Gazzetta Ufficiale del 25 settembre 2025 la legge 23 settembre 2025, n. 132, intitolata «Disposizioni e deleghe al Governo in materia di intelligenza artificiale».

In verità il Senato aveva già approvato il disegno di legge in prima lettura il 20 marzo 2025 (con numero 1146 e titolo «Disposizioni e delega al Governo in materia di intelligenza artificiale»), ma la Camera dei deputati, poi, aveva introdotto delle modificazioni e si è reso così necessario un nuovo vaglio da parte della Camera Alta.

In questa seconda lettura il disegno di legge era stato assegnato, in sede referente, alle commissioni riunite 8ª (Ambiente, transizione ecologica, energia, lavori pubblici, comunicazioni, innovazione tecnologica) e 10ª (Affari sociali, sanità, lavoro pubblico e

[*] Professore associato di Diritto costituzionale, Università degli Studi Niccolò Cusano.

privato, previdenza sociale) e sono stati richiesti i pareri delle commissioni 1ª (Affari costituzionali, affari della Presidenza del Consiglio e dell'Interno, ordinamento generale dello Stato e della Pubblica Amministrazione, editoria, digitalizzazione), 2ª (Giustizia), 3ª (Affari esteri e difesa), 4ª (Politiche dell'Unione europea), 5ª (Programmazione economica, bilancio), 7ª (Cultura e patrimonio culturale, istruzione pubblica, ricerca scientifica, spettacolo e sport) e 9ª (Industria, commercio, turismo, agricoltura e produzione agroalimentare).

Seppure, a legge approvata, possa sembrare non essenziale riportare anche l'elenco delle commissioni chiamate a pronunciarsi in sede consultiva, è apparso utile procedere comunque in tal senso onde rappresentare i diversi e (direi) innumerevoli aspetti che la materia "intelligenza artificiale" attraversa, così da rendere "plasticamente" l'idea della pervasività di tale nuova tecnologia, che ha reso necessaria l'integrazione delle competenze pur cospicue delle commissioni operanti in sede referente. Non è certo un caso, in fondo, che Luciano Floridi inizi così la sua riflessione sull' «Etica dell'intelligenza artificiale»: «Istruzione, affari e industria, viaggi e logistica, banche, vendita al dettaglio e shopping, intrattenimento, welfare e sanità, politica e relazioni sociali, in breve la vita stessa per come la conosciamo oggi è diventata inconcepibile senza la presenza di pratiche, prodotti, servizi e tecnologie digitali»[1]. D'altra parte, la dottrina costituzionalistica già si è preoccupata di definire (per quanto possibile) un «ordine giuridico del digitale»[2], nonché un «ordine giuridico dell'algoritmo»[3], proprio perché

[1] L. FLORIDI, *Etica dell'intelligenza artificiale. Sviluppi, opportunità, sfide*, Edizione italiana a cura di M. Durante, Raffaello Cortina Editore, Milano, 2022, 11.

[2] Vedi T. E. FROSINI, *L'ordine giuridico del digitale*, in «Giur. cost.», 2023, 377 ss.

[3] Vedi A. STERPA (a cura di), *L'ordine giuridico dell'algoritmo*, Editoriale Scientifica, Napoli, 2024.

«l'AI è ormai, e sempre più lo sarà, un elemento essenziale nella vita delle persone, nel sistema di relazioni sociali, che è proprio ciò che il diritto è chiamato a mediare con le sue regole e le sue risorse»[4].

La Relazione di accompagnamento del disegno di legge individua obiettivi davvero ambiziosi: «proteggere i diritti fondamentali, la democrazia, lo Stato di diritto, la sostenibilità ambientale […] promuovendo nel contempo l'innovazione per il benessere dei cittadini».

Non solo; viene precisato che s'intende «operare un bilanciamento tra opportunità e rischi» in vista di un «utilizzo delle nuove tecnologie» volto al «miglioramento delle condizioni di vita dei cittadini e della coesione sociale» e l'individuazione di «soluzioni del rischio fondate su una visione antropocentrica».

Del resto, il Regolamento europeo sull'intelligenza artificiale «è impostato su un'architettura di rischi» derivanti dall'impiego dell'intelligenza artificiale graduati su di una scala che li vede qualificati come «inaccettabili, alti, limitati e minimi»: al crescere del livello di rischio è collegata una maggiore responsabilità o una maggiore intensità dei divieti in capo a coloro che sviluppano o utilizzano sistemi di intelligenza artificiale.

Resta fermo che questo disegno di legge «non si sovrappone» al regolamento europeo, bensì «ne accompagna il quadro regolatorio in quegli spazi propri del diritto interno»[5].

Il Regolamento europeo del 13 giugno 2024 «che stabilisce regole armonizzate sull'intelligenza artificiale», c.d. AI Act, si propone obiettivi di certo non meno ambiziosi di quelli individuati *de*

[4] Così A. D'Aloia, *Il diritto verso "il mondo nuovo". Le sfide dell'Intelligenza Artificiale*, in «BioLaw Journal –Rivista di BioDiritto», 1, 2019, 11.

[5] Vedi *Atti Senato*, XIX Legislatura, Disegno di Legge n. 1146 "Disposizioni e delega al Governo in materia di intelligenza artificiale", 3.

iure condendo sul piano interno e ora positivizzati agli articoli 1 (Finalità e ambito di applicazione) e 3 (Principi generali) della legge n. 132/2025.

Il Considerando n. 1, infatti, così prevede: «Lo scopo del presente regolamento è migliorare il funzionamento del mercato interno istituendo un quadro giuridico uniforme in particolare per quanto riguarda lo sviluppo, l'immissione sul mercato, la messa in servizio e l'uso di sistemi di intelligenza artificiale (sistemi di IA) nell'Unione, in conformità dei valori dell'Unione, promuovere la diffusione di un'intelligenza artificiale (IA) antropocentrica e affidabile, garantendo nel contempo un livello elevato di protezione della salute, della sicurezza e dei diritti fondamentali sanciti dalla Carta dei diritti fondamentali dell'Unione europea («Carta»), compresi la democrazia, lo Stato di diritto e la protezione dell'ambiente, proteggere contro gli effetti nocivi dei sistemi di IA nell'Unione, nonché promuovere l'innovazione».

Obiettivi subito ribaditi al successivo Considerando n. 2: «Il presente regolamento dovrebbe essere applicato conformemente ai valori dell'Unione sanciti dalla Carta agevolando la protezione delle persone fisiche, delle imprese, della democrazia e dello Stato di diritto e la protezione dell'ambiente, promuovendo nel contempo l'innovazione e l'occupazione e rendendo l'Unione un leader nell'adozione di un'IA affidabile».

Non si può non notare come laddove si prescrive l'applicazione del Regolamento in conformità «ai valori dell'Unione sanciti dalla Carta» venga utilizzato il modo condizionale («dovrebbe essere applicato»; «*should be applied*» nel testo inglese; «*devrait être appliqué*» in quello francese): all'orecchio del giurista continentale suona davvero strano.

La scelta del condizionale potrebbe anche ricollegarsi al contesto "preambolare" proprio dei «considerando», ove si delineano

magari scopi da perseguire od anche indirizzi interpretativi in un quadro in un certo senso di "prescrittività attenuata" rispetto all'articolato in senso stretto del Regolamento.

Comunque, sarebbe del tutto illogico concludere che l'uso del condizionale possa lasciar spazio ad un'opzione ermeneutica per cui siano ammissibili ipotesi di applicazione del Regolamento in deroga «ai valori dell'Unione sanciti dalla Carta», soprattutto nel momento in cui si intende definire un quadro comune di tutela da impieghi abusivi dell'intelligenza artificiale. Peraltro, fermo l'uso del modo condizionale anche nelle lingue diverse dall'inglese, va rilevato che nella *English Style Guide* della Commissione Europea si legge: «Positive imperative. *To impose an obligation or a requirement, EU legislation uses* shall […] *Here*, shall *means the same as* must»[6].

È evidente, in ogni modo, come il legislatore eurounitario e quello nazionale siano ben consapevoli dell'impatto sistemico dell'impiego di una tecnologia quale è l'intelligenza artificiale, che può esser fonte di opportunità prima inimmaginabili, ma che con (almeno) egual forza può esser fonte anche di «rischi» per la salute dell'uomo e per la tenuta delle regole fondamentali dell'ordinato vivere comune. Di qui l'esigenza dell'intervento normativo nell'ottica, per quanto possibile, di massimizzare le opportunità e di contenere i rischi; intento, questo, che diventa ancor più pressante se ad esser esposte ai rischi sono le persone in condizione di (particolare) vulnerabilità[7] o se si pone mente al fatto che «l'IA può creare

[6] EUROPEAN COMMISSION, *English Style Guide*. A handbook for authors and translators in the European Commission, Eighth edition: January 2016, Last updated: August 2025, 61. Per l'uso di *shall* in ambito contrattuale vedi C. C. BUSK, *Using* Shall *or* Will *to Create Obligations in Business Contracts. Once More unto the Breach*, in «Michigan Bar Journal», October 2017, 50 ss.

[7] Sul punto vedi *infra* § 5.

vulnerabilità nuove, inducendo meccanismi di esclusione sociale»[8], fermo che – si è osservato – un eccesso di regolazione porta con sé delle controindicazioni anche in questa materia[9].

Della rilevanza giuridica della vulnerabilità è d'altronde ormai difficile dubitare, attesi l'espresso riferimento fatto ad essa nel Regolamento europeo, l'emersione del concetto nella legislazione interna[10] e nella giurisprudenza costituzionale[11], il contributo dato

[8] Vedi M. TOMASI, L. BUSATTA, M. FASAN, C. NARDOCCI, S. PENASA, S. SULMICELLI, *Vulnerabilità e Intelligenza Artificiale*, in «BioLaw Journal – Rivista di BioDiritto», Special Issue 1, 2024, 1.

[9] Sul punto vedi S. PALEARI, F. RESTA, *Regole, ma l'AI non va frenata*, in *Corriere della Sera*, 6 agosto 2025. Sul problema di "come" regolare una materia così particolare (per certi versi insidiosa) quale è l'intelligenza artificiale vedi già le riflessioni di A. SIMONCINI, *Verso la regolamentazione della Intelligenza Artificiale. Dimensioni e governo*, in «BioLaw Journal – Rivista di BioDiritto», 2, 2021, 411 ss., nonché, più di recente, F. CIRILLO, *Un diritto per l'intelligenza artificiale:* Quis, Quid, Quando, Ubi, Cur*?*, in «Democrazia e Diritti Sociali», 2, 2024, 199 ss.; G. DE MINICO, *Le fonti del diritto: un argine all'intelligenza artificiale?*, in «Rivista AIC», 3, 2025, 76 ss. Su profili non solo di merito, ma anche di metodo della disciplina introdotta dal Regolamento europeo vedi: gli interventi nel Forum *Il Regolamento sull'intelligenza artificiale*, con *Introduzione* di M. TROISI, in «Gruppo di Pisa. Dibattito aperto sul Diritto e la Giustizia Costituzionale», 2, 2024, 199 ss.; F. DONATI, *La protezione dei diritti fondamentali nel regolamento sull'intelligenza artificiale*, in «Rivista AIC», 1, 2025, 1 ss.; da ultimo: gli interventi nel *Forum* introdotto da C. CASONATO, *AI Act: un dialogo multidisciplinare. Introduzione*, in «BioLaw Journal – Rivista di BioDiritto», 3, 2025, 227 ss. Per la regolazione della IA sul piano internazionale vedi F. ZORZI GIUSTINIANI, *La Convenzione quadro sull'intelligenza artificiale e i diritti umani, la democrazia e lo Stato di diritto e l'*International Cyberspace and Digital Policy Strategy *statunitense*, in «Nomos. Le attualità nel diritto», 2, 2024. In fondo poiché «il diritto costituzionale è direttamente investito dallo sviluppo delle nuove tecnologie» ci si è interrogati anche sul problema della regolazione del Metaverso: R. BIFULCO, *Riverberi costituzionali del Metaverso*, in «MediaLaws», 3, 2023, 41 ss.

[10] Sul punto vedi M. LUCIANI, *Le persone vulnerabili e la Costituzione*, 22 aprile 2022 (*https://www.cortecostituzionale.it/documenti/convegni_seminari/roma_2022-persone_vulnerabili_-_massimo_luciani_20220503170920.pdf*).

alla determinazione dei suoi contorni giuridici dalla Corte di giustizia dell'Unione europea[12] e dalla Corte europea dei diritti dell'uomo[13], l'attenzione riservata a tale categoria dalla dottrina (anche quando investita di ruoli istituzionali)[14].

2. Vulnerabilità e disabilità nel Regolamento europeo

Giova ora procedere ad una disamina più specifica del contenuto del Regolamento europeo del 13 giugno 2024 onde delineare il complessivo quadro regolatorio eurounitario di tutela delle persone con disabilità a fronte di impieghi di sistemi di intelligenza artificiale.

Il Regolamento annovera tra le condizioni di vulnerabilità la posizione delle persone con disabilità, cui fa intanto riferimento in ben undici «considerando» a proposito di: tecniche di manipolazione del processo decisionale (n. 29); identificazione biometrica remota (n. 32); «impatto negativo» sui diritti protetti dalla Carta dei diritti fondamentali dell'Unione europea (n. 48); dati biometrici (n. 54); istruzione (n. 56); lavoro (n. 57); servizi pubblici e priva-

[11] Vedi, da ultimo, Corte cost., 29 dicembre 2025, n. 204, commentata da A. RUGGERI, *La "regionalizzazione" della disciplina del fine-vita: un bicchiere mezzo pieno per le fonti di autonomia e mezzo vuoto per la Costituzione (a prima lettura di Corte cost. n. 204 del 2025)*, in «Consulta Online», Editoriale 2026/I (3 gennaio 2026), nonché Corte cost., 23 gennaio 2025, n. 3, su cui vedi *infra* § 4.

[12] Vedi, se si vuole, F. GIRELLI, *Il contributo della corte europea dei diritti dell'uomo all'elaborazione di un diritto a una morte dignitosa*, in «Diritto e Società», 2, 2023, 345 ss. e in T. CERRUTI (a cura di), *L'elaborazione di un diritto a una morte dignitosa nell'esperienza europea*, Atti del Seminario Annuale dell'Associazione "Gruppo di Pisa", Università degli Studi di Torino, 2 dicembre 2022, Editoriale Scientifica, Napoli, 2023, 42 ss.

[13] Vedi *infra* § 5.

[14] Vedi *supra* note 8 e 10 e *infra* note 15, 40, 67-71 e 75.

ti (n. 58); rispetto della Convenzione delle Nazioni Unite sui diritti delle persone con disabilità e piena considerazione «delle loro dignità e diversità intrinseche» (n. 80); sistemi di IA destinati alla interazione con le persone o a generare contenuti (n. 132); «soluzioni basate sull'IA per aumentare l'accessibilità per le persone con disabilità» intese come «risultati vantaggiosi dal punto di vista sociale e ambientale» della ricerca in tema di IA (n. 142); promozione di un' «IA etica ed affidabile» (n. 165).

Ma vediamo come poi nell'articolato del Regolamento, quindi non più nel contesto "preambolare", viene tutelata la peculiare condizione di vulnerabilità riconosciuta alle persone con disabilità.

Subito viene in rilievo l'articolo 5 che tra le «Pratiche di IA vietate» annovera [alla lettera b) del comma 1] «l'immissione sul mercato, la messa in servizio o l'uso di un sistema di IA che sfrutta le vulnerabilità di una persona fisica o di uno specifico gruppo di persone, dovute all'età, alla disabilità o a una specifica situazione sociale o economica, con l'obiettivo o l'effetto di distorcere materialmente il comportamento di tale persona o di una persona che appartiene a tale gruppo in un modo che provochi o possa ragionevolmente provocare a tale persona o a un'altra persona un danno significativo»[15].

[15] Si è osservato che «sarebbe stato auspicabile anche il riconoscimento del fatto che la vulnerabilità può derivare (oltre che da caratteristiche personali o situazionali) anche dall'interazione tra le caratteristiche particolari di un individuo e un sistema di IA», così C. M. CASCIONE, *Sub Articolo 5 – Pratiche di IA vietate, Commento comma 1 lett. b)*, in A. MANTELERO, G. RESTA, G. M. RICCIO (a cura di), *Intelligenza artificiale. Commentario*, I Edizione, Wolters Kluwer, Milano, 2025, 142. Sulle potenzialità ermeneutiche offerte da questa disposizione in vista di una ricostruzione soddisfacente del concetto di vulnerabilità nell'ambito del Regolamento europeo vedi, in ogni modo, F. GALLI, C. NOVELLI, *The Many Meanings of Vulnerability in the AI Act and the One Missing*, in «BioLaw Journal – Rivista di BioDiritto», Special Issue 1, 2024, 53 ss.

Il divieto, quindi, dato il tenore testuale della previsione («con l'obiettivo o l'effetto»), opera a prescindere dall'intento di chi implementa queste «Pratiche» ed è volto, pertanto, a contrastare l'effetto distorsivo in quanto tale comunque si produca.

L'articolo 60, poi, rubricato «Prove di sistemi di IA ad alto rischio in condizioni reali al di fuori degli spazi di sperimentazione normativa per l'IA», dopo aver riaffermato i divieti sanciti dall'articolo 5, al comma 4 prescrive le condizioni da soddisfare da parte dei «fornitori o potenziali fornitori» di sistemi di IA ad alto rischio per poter effettuare prove «in condizioni reali». In particolare, la lettera g) del citato comma 4 stabilisce che debbano essere «adeguatamente protetti» coloro che appartengono a «gruppi vulnerabili a causa della loro età o disabilità».

Inoltre, l'articolo 95, comma 2, prevede che l'ufficio per l'IA[16] e gli Stati membri agevolino l'elaborazione di codici di condotta relativi all'applicazione volontaria di requisiti specifici a tutti i sistemi di IA. Questa attività, indirizzata anche ai deployer[17], deve contemplare «obiettivi chiari e indicatori chiave di prestazione volti a misurare il conseguimento di tali obiettivi» e, in tal senso, «a titolo puramente esemplificativo» sono indicati, altresì, alcuni

[16] L'*European Artificial Intelligence Office* è stato istituito con la Decisione della Commissione Europea del 24 gennaio 2024. Resta fermo che l'art. 3 del Regolamento del 13 giugno 2024, rubricato «Definizioni», al n. 47) così stabilisce: «"ufficio per l'IA": la funzione della Commissione volta a contribuire all'attuazione, al monitoraggio e alla supervisione dei sistemi di IA e dei modelli di IA per finalità generali, e della governance dell'IA prevista dalla decisione della Commissione del 24 gennaio 2024. I riferimenti all'ufficio per l'IA contenuti nel presente regolamento si intendono fatti alla Commissione».

[17] L'art. 3, n. 4), del Regolamento del 13 giugno 2024 così definisce il deployer: «una persona fisica o giuridica, un'autorità pubblica, un'agenzia o un altro organismo che utilizza un sistema di IA sotto la propria autorità, tranne nel caso in cui il sistema di IA sia utilizzato nel corso di un'attività personale non professionale».

«elementi» da necessariamente considerare. Tra questi ultimi va in particolare segnalato quello previsto alla lettera e): «la valutazione e la prevenzione dell'impatto negativo dei sistemi di IA sulle persone vulnerabili o sui gruppi di persone vulnerabili, anche per quanto riguarda l'accessibilità per le persone con disabilità, nonché sulla parità di genere».

Visto quanto stabilito già nei Considerando n. 1 e n. 2, ossia, in particolare, che con il Regolamento si intende intervenire «in conformità dei valori dell'Unione», garantire «i diritti fondamentali sanciti dalla Carta dei diritti fondamentali dell'Unione europea» e che il Regolamento medesimo «dovrebbe essere applicato conformemente ai valori dell'Unione sanciti dalla Carta»[18], non ci si può adesso esimere dal richiamare quanto previsto in modo espresso appunto dalla Carta dei diritti fondamentali dell'Unione europea circa le persone con disabilità.

L'articolo 26, rubricato «Inserimento delle persone con disabilità», collocato nel Capo III, intitolato «Uguaglianza», così prescrive: «L'Unione riconosce e rispetta il diritto delle persone con disabilità di beneficiare di misure intese a garantirne l'autonomia, l'inserimento sociale e professionale e la partecipazione alla vita della comunità».

La disposizione va interpretata naturalmente in chiave sistematica; basti, in ogni modo, qui ricordare almeno altre due disposizioni, che, peraltro, aprono il Capo III della Carta: l'articolo 20 che stabilisce che «*Tutte le persone* sono uguali davanti alla legge» (corsivo mio) e l'articolo 21 che sancisce il divieto di discriminazione in ragione della disabilità.

[18] Vedi *supra* § 1.

3. La legge 23 settembre 2025, n. 132

Occorre ora verificare come a livello interno si è provveduto a raccordarsi sul piano positivo con tale quadro regolatorio eurounitario.

Il legislatore ha inteso, s'è detto, "accompagnare" tale quadro regolatorio eurounitario tramite la disciplina ora recata dalla recentissima legge 23 settembre 2025, n. 132.

L'attenzione alle persone con disabilità emerge già nei «Princìpi generali» individuati nell'art. 3, giacché al comma 7 è stabilito che la «presente legge garantisce alle persone con disabilità il pieno accesso ai sistemi di intelligenza artificiale e alle relative funzionalità o estensioni, su base di uguaglianza e senza alcuna forma di discriminazione e di pregiudizio, in conformità alle disposizioni della Convenzione delle Nazioni Unite sui diritti delle persone con disabilità».

Al comma 4, poi, dell'art. 7, rubricato «Uso dell'intelligenza artificiale in ambito sanitario e di disabilità»[19], è previsto che sempre la legge in parola promuova «lo sviluppo, lo studio e la diffusione di sistemi di intelligenza artificiale che migliorano le condizioni di vita delle persone con disabilità, agevolano l'accessibilità, la mobilità indipendente e l'autonomia, la sicurezza e i processi di inclusione sociale delle medesime persone anche ai fini dell'elabo-

[19] All'indomani dell'approvazione definitiva da parte del Senato del disegno di legge n. 1146-B i primi commentatori hanno osservato: «L'art. 7 stabilisce **principi condivisibili** per l'uso dell'IA in sanità, come quelli di non discriminazione, informazione e trasparenza, inclusione sociale, imprescindibilità della decisione umana, accuratezza e sicurezza dei sistemi. Tuttavia, sarebbe stato auspicabile un migliore coordinamento con l'AI Act per integrare prescrizioni più specifiche sulla qualità dei dati, tra cui la preferenza per i dati sintetici e anonimizzati», così G. RESTA, G. M. RICCIO, *Il ddl AI è legge: primi spunti di riflessione*, in «Il Quotidiano Giuridico», 18 settembre 2025.

razione del progetto di vita di cui all'articolo 2, comma 1, lettera *n)*, del decreto legislativo 3 maggio 2024, n. 62». Tale decreto, intitolato «Definizione della condizione di disabilità, della valutazione di base, di accomodamento ragionevole, della valutazione multidimensionale per l'elaborazione e attuazione del progetto di vita individuale personalizzato e partecipato», è entrato in vigore il 30 giugno 2024 ed è uno dei decreti legislativi adottati in attuazione della legge 22 dicembre 2021, n. 227 recante «Delega al Governo in materia di disabilità»[20].

L'art. 22 (Misure di sostegno ai giovani e allo sport) reca una disposizione, il comma 3, che interessa direttamente la posizione delle persone con disabilità: «Lo Stato favorisce l'accessibilità ai sistemi di intelligenza artificiale per il miglioramento del benessere psicofisico attraverso l'attività sportiva, anche ai fini dello sviluppo di soluzioni innovative finalizzate a una maggiore inclusione in ambito sportivo delle persone con disabilità. Nel rispetto dei princìpi generali di cui alla presente legge, i sistemi di intelligenza artificiale possono essere utilizzati anche per l'organizzazione delle attività sportive»; ed un'altra, il comma 2, che concerne «le studentesse e gli studenti ad alto potenziale cognitivo». Tale ultima previsione, seppure di tenore testuale non propriamente limpido, va comunque valorizzata proprio perché rivela attenzione per questa

[20] Vedi a riguardo E. VIVALDI (a cura di), *Il progetto di vita della persona con disabilità. Dal PNRR al decreto legislativo n. 62/2024*, Pisa University Press, Pisa, 2025; G. ARCONZO, *L'incerto avvio delle riforme in tema di disabilità previste dalla legge n. 227 del 2021*, in «Federalismi.it», 16, 2025, 15 ss. Sul decreto legislativo n. 62/2024 vedi anche F. MASCI, *Persona, scienza e linguaggio: l'inattuazione del programma costituzionale in materia di disabilità e dei relativi livelli essenziali delle prestazioni*, in «Corti supreme e salute», 2, 2024, 635 ss. Sulla centralità del «progetto di vita, individuale, personalizzato e partecipato» già nell'impianto della legge delega n. 227/2021 vedi P. ADDIS, *La legge 227/2021: questioni definitorie e diritti*, in «PasSaggi Costituzionali», 1, 2023, 253 ss.

categoria di studenti non sempre debitamente seguiti[21]. Paradossalmente la loro specifica (vantaggiosa) condizione può tradursi in una forma di vulnerabilità, nel momento in cui non vengono adottate le necessarie misure idonee a sostenerla affinché questo «alto potenziale» possa correttamente esprimersi[22]; del resto, è noto come la condizione stessa di disabilità possa essere acuita, se non prodotta, dal contesto in cui la persona è inserita[23].

L'opera di inclusione per essere autentica non può tollerare la marginalizzazione di chi comunque è portatore di un particolare bisogno. Non solo; per essere efficace va supportata con adeguate risorse e sotto questo aspetto emerge il profilo più critico di questa legge ossia la «Clausola di invarianza finanziaria» di cui all'art. 27, che consente un impegno aggiuntivo di spesa per la sola «Applicazione sperimentale dell'intelligenza artificiale ai servizi forniti dal

[21] Art. 22, comma 2, legge n. 132/2025: «Nel piano didattico personalizzato (PDP) adottato dall'istituzione scolastica secondaria di secondo grado nell'ambito della propria autonomia, per le studentesse e gli studenti ad alto potenziale cognitivo possono essere inserite attività volte all'acquisizione di ulteriori competenze attraverso esperienze di apprendimento presso le istituzioni della formazione superiore, in deroga a quanto previsto dall'articolo 6 del regolamento di cui al decreto del Ministro dell'istruzione, dell'università e della ricerca 22 ottobre 2004, n. 270, in materia di previo possesso del titolo, in coerenza con il profilo in uscita dell'indirizzo di studio. I crediti formativi acquisiti nell'ambito delle attività di cui al primo periodo sono valutati nell'ambito dei percorsi formativi della formazione superiore intrapresi dopo il conseguimento del titolo di cui al medesimo primo periodo».

[22] Vedi L. MILAN, *Chi sono gli studenti gifted e come riconoscerli?*, 30 luglio 2025 (*https://www.erickson.it/it/mondo-erickson/chi-sono-gli-studenti-gifted-e-come-riconoscerli*). Della posizione di tale categoria di studenti si è interessata di recente anche la giurisprudenza amministrativa, vedi TAR Veneto, 8 settembre 2025, n. 1530.

[23] Sul punto, se si vuole, vedi F. GIRELLI, *La disabilità e il corpo nella dimensione costituzionale*, in «Consulta Online», II, 2021, 608-609.

Ministero degli affari esteri e della cooperazione internazionale» come previsto dall'art. 21.

Nemmeno, onestamente, stupisce più di tanto questa scelta di escludere «nuovi o maggiori oneri a carico della finanza pubblica», in quanto ormai, purtroppo, previsioni di tal fatta sono divenute una sorta di clausola di stile del legislatore.

Nondimeno resta il "contenuto normativo" di questa legge, che, anzi, non manca, nell'introdurre (all'art. 26) la nuova fattispecie di reato «Illecita diffusione di contenuti generati o alterati con sistemi di intelligenza artificiale», perseguibile a querela della persona offesa, di stabilire la procedibilità d'ufficio anche per questa ipotesi, qualora il fatto sia «commesso nei confronti di persona incapace, per età o per infermità».

Quel che pare emergere dalla lettura complessiva del testo della legge, dunque, è la fissazione di un divieto generalizzato di arrecare danno alle persone con disabilità nel momento in cui si impiegano sistemi di intelligenza artificiale; divieto, questo, che poi trova il suo risvolto, come dire, "in positivo" nelle previsioni che stabiliscono l'utilizzo dell'intelligenza artificiale al solo scopo di supportare le persone con disabilità e di migliorarne semmai le condizioni di vita. Insomma, l'altra faccia della medaglia di tale divieto istituito per l'ambito tecnologico più avanzato è un vero e proprio *dovere di solidarietà digitale*: il trascorrere del tempo e l'evoluzione (velocissima) della tecnica non attenuano certo la forza prescrittiva dell'art. 2 Cost., che trova ora nuove declinazioni e risponde ad esigenze magari impensabili anche fino a poco tempo fa. In fondo, il testo elaborato dal Costituente «proprio perché pensato per durare *aere perennius*, esige dall'interprete di essere sondato – e rispettato – in tutte le sue virtualità di significato»[24].

[24] Così M. LUCIANI, *Interpretazione costituzionale e testo della Costituzione. Osservazioni liminari*, in G. AZZARITI (a cura di), *Interpretazione costituzionale*,

Soprattutto adesso, quindi, nell'«età della tecnica», appare significativo riaffermare il richiesto «adempimento dei doveri inderogabili di solidarietà», in quanto il diritto vive la stagione del «tecnodiritto» qualificata dal «*rapporto della potenza giuridica con altre potenze*», *in primis* con la tecnica, che è «potenza che *usa il mondo*, e perciò lo calcola, lo governa, lo manipola»[25].

4. La necessità di un "cyber-biodiritto"

Il Regolamento del Parlamento europeo e del Consiglio del 13 giugno 2024 e la legge n. 132/2025 si pongono in effetti, almeno per quanto qui interessa, su di una linea di convergenza, rinsaldata espressamente sul piano positivo dall' art. 1 della legge, che al secondo comma impone di interpretare e applicare le disposizioni in essa contenute «conformemente al regolamento (UE) 2024/1689 del Parlamento europeo e del Consiglio, del 13 giugno 2024».

Col Regolamento, infatti, si è inteso approntare un quadro normativo volto ad impedire che l'impiego dell'IA possa ledere le condizioni di vulnerabilità o comunque produrre effetti discriminatori; e fondamentale, in questo senso, è il sopra ricordato Considerando n. 80, che sancisce come si debba «tenere pienamente conto delle […] dignità e diversità intrinseche» delle persone con disabilità.

Dalla lettura combinata del Regolamento europeo e della legge n. 132/2025, insomma, si evince l'intento di fissare le condizioni per assicurare l'inclusione delle persone con disabilità anche in un contesto ove pervasivo sia l'utilizzo di nuove tecnologie, in linea,

Giappichelli, Torino, 2007, 49.

[25] Vedi N. IRTI, *Tecno-diritto*, in ID., *Il diritto nell'età della tecnica*, Editoriale Scientifica, Napoli, 2007, 13-14.

peraltro, con le «Priorità» 2 (Accesso e accessibilità) e 5 (Promozione delle nuove tecnologie) della *Carta di Solfagnano*, sottoscritta dai Ministri del G7 per le politiche relative alle persone con disabilità e all'inclusione il 16 ottobre 2024[26]. Ora, presupposto per raggiungere tale obiettivo di *inclusione artificiale*, appunto perché realizzata tramite il corretto e utile impiego di strumenti di IA, pare essere l'adempimento di quello che, considerata la materia, abbiamo definito "dovere di solidarietà digitale", un dovere generalizzato che, agevolmente, almeno per quanto concerne l'ordinamento interno, può radicarsi negli artt. 2 e 3 della Costituzione[27]; del resto, il cuore della tutela costituzionale delle persone con disabilità è proprio la garanzia del "meta-diritto" alla socializzazione[28].

Il punto è che ci avviciniamo sempre più velocemente all' "ultima frontiera"; anzi, a onor del vero, viviamo già nell'era in cui le conoscenze tecnologiche sono così avanzate che l'ambizione umana pare non avere confini: arriva a lambire l'immortalità, a consi-

[26] Vedi la *Carta di Solfagnano* adottata nel G7 *Inclusione e disabilità* nell'ambito del G7 Italia 2024 (*https://www.disabilita.governo.it /media/wztmr4a1/g7-inclusion-and-disability_ solfagnano-charter_ita.pdf*).

[27] Fermo che l'adempimento dei doveri di solidarietà anche da parte delle stesse persone con disabilità «realizza una effettiva parità di *chances* nella realizzazione e nello sviluppo della propria soggettività», così E. VIVALDI, *L'esercizio dei doveri di solidarietà da parte della persona con disabilità*, in «Osservatorio AIC», 4, 2019, 155.

[28] A riguardo sia consentito rinviare a C. COLAPIETRO, F. GIRELLI, *Persone con disabilità e Costituzione*. Presentazione di F. Modugno, Editoriale Scientifica, Napoli, 2020, *passim*; nonché vedi C. COLAPIETRO, *La persona con disabilità nella Costituzione italiana e nel diritto sovranazionale*, in «PasSaggi Costituzionali», 1, 2023, 104 ss. Sui perduranti problemi di istituzionalizzazione e segregazione delle persone con disabilità ed i relativi strumenti di prevenzione e contrasto vedi ora C. TARANTINO (a cura di), *Il soggiorno obbligato. La disabilità fra dispositivi di incapacitazione e strategie di emancipazione*, il Mulino, Bologna, 2024.

derare sul serio di poter «trasformare *Homo sapiens* in *Homo Deus*»[29].

La peculiarità di possibili impieghi dell'intelligenza artificiale, del resto, sta nel fatto che potrebbero incidere su quanto qualifica la nostra umanità (o quantomeno ciò che noi riteniamo tale).

Non siamo di fronte, cioè, ad un mero ausilio che "all'esterno di noi" in qualche modo potenzia le nostre capacità come un qualsiasi strumento, di cui l'uomo nel corso della storia si è via via dotato a partire dall'utensile più rudimentale fino all'apparecchiatura più raffinata. Ci stiamo confrontando, in realtà, con una tecnologia potenzialmente idonea ad impattare su quanto noi s'intende per "umano", già adesso capace, grazie alla sua «potenza cibernetica», di sostituire «l'autonomia» (dell'uomo) con «l'automazione» (della macchina), determinando preoccupanti conseguenze sul contenuto e la sostanza delle libertà fondamentali[30].

I tentativi di regolazione dell'intelligenza artificiale sia a livello interno sia a livello eurounitario, quindi, sono da considerarsi interventi normativi ascrivibili al biodiritto, che, pur nella «persistenza di un pluralismo valoriale per questioni eticamente sensibili

[29] Vedi Y. N. HARARI, *Homo Deus. Breve storia del futuro* [2015], traduzione dall'inglese di M. Piani, Giunti Editore/Bompiani, Firenze-Milano, 2019, 31. E tenere a mente il «celebre mito del "volo di Icaro"» aiuterebbe a metterci in guardia sui rischi di tale ambizione: L. TRUCCO, *Natura e sentimento nel diritto*, Mimesis, Milano, 2024, 183.

[30] Vedi A. SIMONCINI, *L'algoritmo incostituzionale: intelligenza artificiale e il futuro delle libertà*, in «BioLaw Journal –Rivista di BioDiritto», 1, 2019, 89. Del resto, anche il monito di Papa Leone XIV è ben chiaro: «Nel tempo dell'intelligenza artificiale, in cui la dignità umana rischia di essere oscurata da nuove forme di disumanizzazione, abbiamo il dovere urgente di restare profondamente umani» (*Lettera Enciclica* Magnifica Humanitas *del Santo Padre Leone XIV sulla custodia della persona umana nel tempo dell'intelligenza artificiale*, § 15: *https://www.vatican.va/content/leo-xiv/it/encyclicals/documents/20260515-magnifica-humanitas.html*).

e ancora divisive» tra i singoli Stati, costituiscono comunque un esempio di quella che è stata definita «collaborazione di tipo *verticale*»[31]. E non è un caso che la legge n. 132/2025 ed il Regolamento europeo facciano riferimento ad una prospettiva «antropocentrica»[32]: emerge qui la consapevolezza da parte del legislatore circa il rischio di perdita della nostra umanità o comunque di una "contaminazione" della stessa in misura tale da non renderla più riconoscibile, da non renderci più riconoscibili quali esseri umani. Nell'attuale contesto, pertanto, il «diritto non può starsene come un curioso spettatore, né delegare ad altre potenze la guida degli uomini. I tempi esigono una presa di posizione»[33].

Di qui la necessità di un "cyber-biodiritto", proprio perché la «tecnologia è sempre più integrata con la vita della persona e questa integrazione non è realisticamente arrestabile, né dall'altra parte ciò sarebbe auspicabile, visto il valore altissimo per la stessa qua-

[31] Vedi L. CHIEFFI, *Brevi note su Europa e il Biodiritto*, in «BioLaw Journal – Rivista di BioDiritto», 1, 2023, 54.

[32] Vedi art. 1, comma 1, legge n. 132/2025: «La presente legge reca principi in materia di ricerca, sperimentazione, sviluppo, adozione e applicazione di sistemi e di modelli di intelligenza artificiale. Promuove un utilizzo corretto, trasparente e responsabile, in una dimensione antropocentrica, dell'intelligenza artificiale, volto a coglierne le opportunità. // Garantisce la vigilanza sui rischi economici e sociali e sull'impatto sui diritti fondamentali dell'intelligenza artificiale»; vedi anche *supra* § 1 e il Considerando n. 6 del Regolamento europeo: «In considerazione dell'impatto significativo che l'IA può avere sulla società e della necessità di creare maggiore fiducia, è essenziale che l'IA e il suo quadro normativo siano sviluppati conformemente ai valori dell'Unione sanciti dall'articolo 2 del trattato sull'Unione europea (TUE), ai diritti e alle libertà fondamentali sanciti dai trattati e, conformemente all'articolo 6 TUE, alla Carta. Come prerequisito, l'IA dovrebbe essere una tecnologia antropocentrica. Dovrebbe fungere da strumento per le persone, con il fine ultimo di migliorare il benessere degli esseri umani».

[33] Così N. IRTI, *Bio-diritto*, in ID., *Il diritto nell'età della tecnica*, Editoriale Scientifica, Napoli, 2007, 45.

lità della vita che alcune nuove scoperte hanno»[34]. Il problema di fondo resta questo: non tutto ciò che è tecnicamente possibile è consentito o è opportuno realizzare, come, del resto, il progresso della scienza medica testimonia quotidianamente.

Tutto ciò è evidente in specie nei casi estremi[35]; non sono certamente l'impiego dell'intelligenza artificiale per facilitare la mobilità urbana[36], i sistemi di domotica o il cosiddetto "bastone intelligente" per le persone non vedenti[37] a costringerci ad interrogarci su tali questioni capitali. Questi sono esempi di supporto senza dubbio virtuoso, assicurato grazie all'intelligenza artificiale, così come lo sono anche i suoi possibili impieghi in ambito scolastico o lavorativo. In forza della sua «capacità di personalizzare» l'intervento educativo l'IA può operare non solo come «supporto», ma come «un vero promotore di inclusione». In realtà, «in piena coe-

[34] Vedi A. SIMONCINI, *L'algoritmo incostituzionale: intelligenza artificiale e il futuro delle libertà*, cit., 88.

[35] Ad avviso di M. LUCIANI [dicembre 2023], *Libertà di ricerca e intelligenza artificiale*, in ASSOCIAZIONE ITALIANA DEI COSTITUZIONALISTI, *Lettere AIC. 2021-2024*, Editoriale Scientifica, Napoli, 2024, 289-290, la regolazione della ricerca scientifica dovrebbe interessare appunto i suoi «*confini estremi*». Il contributo di Massimo Luciani è stato pubblicato dapprima sul sito web dell'AIC, il 17 dicembre 2023, con il titolo *La sfida dell'intelligenza artificiale* (*https://www.associazionedeicostituzionalisti.it/images/laLettera/LaLetteraAIC_12_2023_Contributo_introduttivo.pdf).*

[36] Sull'evoluzione della «concezione di città *smart*» vedi A. PAPA, Smart city *e inclusione (non solo) digitale*, in «Nomos. Le attualità nel diritto», 3, 2024.

[37] Vedi *Il bastone per non vedenti che usa l'intelligenza artificiale*, 29 gennaio 2025 (*https://stream24.ilsole24ore.com/video/tecnologia/il-bastone-non-vedenti-che-usa-intelligenza-artificiale/AG6EPpcC?refresh_ce=1*). Per un quadro sulle nuove tecnologie a supporto delle persone con disabilità vedi A. PEPINO, G. SICIGNANO, E. VALLEFUOCO, *Le tecnologie della comunicazione a supporto delle persone con disabilità, stato dell'arte e prospettive future*, in «PasSaggi Costituzionali», 1, 2023, 399 ss.

renza» con il sopra citato decreto legislativo n. 62/2024[38], sono oggi disponibili tecnologie che «consentono alla persona con disabilità di esprimersi con il proprio funzionamento e partecipare attivamente al processo di co-progettazione, valorizzandone il suo ruolo di protagonista nella costruzione del proprio progetto di vita individuale, personalizzato e partecipato»[39]. Anzi, proprio di recente la Corte costituzionale ha affermato come le nuove tecnologie debbano essere poste al servizio delle vulnerabilità (nel caso di specie per favorire la partecipazione democratica delle persone con disabilità)[40], atteso che «la dignità umana è compromessa ogni volta in cui è lo stesso ordinamento giuridico che trasforma, in forza di un suo divieto o di una sua previsione, in inabile e bisognosa di assistenza una persona che, invece, sarebbe in grado, con propri mezzi, di provvedere a compiere una determinata attività»[41].

[38] Vedi *supra* § 3.

[39] Vedi L. FENZA, *IA, disabilità e progetto di vita: nuove prospettive per l'analisi dei bisogni, la quotidianità e la realizzazione del progetto di vita della persona con disabilità*, in «Italian Journal of Special Education for Inclusion», 1, 2025, 319 ss. Sulle possibilità di utilizzo in ambito lavorativo in particolare vedi F. LAMBERTI, *Inclusione lavorativa delle persone con disabilità: le soluzioni basate sull'intelligenza artificiale* come sostegni necessari *e* accomodamenti ragionevoli, in «AmbienteDiritto.it», 3, 2025. Purtroppo la cronaca già ci ha consegnato impieghi dell'IA fortemente lesivi della dignità delle persone con disabilità: E. NICOLOSI, *Video erotici e scritte come "Merito amore anche io", ma dietro il Disability Porn c'è una rete di truffatori*, 10 Luglio 2025, Aggiornato il 25 Agosto 2025 (*https://www.alfemminile.com/culture/filtro-sindrome-di-down-sui-social-e-problematico/?ref=RHRM-BG-P12-S1-T1-fem*).

[40] Sulla «partecipazione come forma di inclusione» vedi V. LORUBBIO, *Partecipazione, Diritti umani e condizioni di vulnerabilità: "virtù e vizi" del diritto internazionale*, in «Nomos. Le attualità nel diritto», 3, 2024, 15 ss.

[41] Vedi Corte cost., 23 gennaio 2025, n. 3, n. 4.2 del *Considerato in diritto*, decisione che (giustamente) ha attratto l'attenzione di diversi commentatori: M. MANCINI, *Un'altra rilevante pronuncia della Corte in materia elettorale, tra orientamenti già maturati, profili innovativi e "non detto"*, in «Giur. cost.», 2025, 12 ss.; D. PICCIONE, *I diritti politici della persona con disabilità tra differenziazione ed*

Sono allora *altre* le ipotesi da prendere in considerazione, ovverosia quelle che reclamano tutela dal biodiritto poiché in esse la «cura della salute dell'essere umano» si pone come problema proprio a cagione del «progresso scientifico e tecnologico»[42].

Pensiamo ad una persona con una disabilità mentale grave dovuta al decorso di una malattia degenerativa. Se fosse possibile con un sistema di intelligenza artificiale potenziare enormemente le sue capacità cognitive, sarebbe forse lecito pensare che grazie alla tecnologia si stia in qualche modo restituendo questa persona a sé stessa.

Pensiamo, invece, al caso di una persona *nata* con una grave disabilità mentale. Se trovasse anche qui applicazione quella stessa

egualitarismo, *ibidem*, 23 ss.; A. FERRARA, *L'illegittimità della disposizione richiamata e non di quella che opera il rinvio normativo: un mero errore materiale?*, *ibidem*, 30 ss.; L. TRUCCO, *La «possibilità» di sottoscrizione digitale delle candidature: dall'amministrazione del Lazio alla libertà della persona*, in «Le Regioni», 3, 2025, 665 ss.; G. ARCONZO, *Sull'incostituzionalità di previsioni che non consentono l'esercizio autonomo dei diritti delle persone con disabilità. Nota a Corte costituzionale, sentenza n. 3 del 2025*, in «Osservatorio AIC», 4, 2025, 194 ss.; F. VIVALDELLI, *Disabilità, diritti politici e democrazia digitale. Considerazioni a margine della sent. n. 3/2025 della corte costituzionale*, in «Gruppo di Pisa. Dibattito aperto sul Diritto e la Giustizia Costituzionale», 1, 2025, 157 ss.; C. LOTTA, *Verso la digitalizzazione del procedimento elettorale preparatorio? Spunti a partire da C. cost. n. 3 del 2025*, in «Nomos. Le attualità nel diritto», 1, 2025; M. GIANNELLI, *L'inclusione attraverso l'innovazione digitale. Dalla Corte costituzionale una "spinta gentile" verso un ripensamento delle regole del procedimento elettorale preparatorio?*, in «lecostituzionaliste.it», 16 aprile 2025; F. SANCHINI, *La Corte costituzionale alle prese con le istanze di "modernizzazione" del procedimento elettorale preparatorio. Nota alla sentenza n. 3 del 2025*, in «Federalismi.it», 28, 2025, 259 ss. Metteva già in luce come per «l'effettiva realizzazione della persona con disabilità» vadano sfruttate «anche le potenzialità delle nuove tecnologie» V. PUPO, *La progressiva attuazione del principio di accessibilità delle persone con disabilità*, in «Rivista AIC», 4, 2023, 95.

[42] Cfr. S. CANESTRARI, *Biodiritto (diritto penale)*, in *Enciclopedia del Diritto*, Annali VIII, Giuffrè, Milano, 2015, 99.

ipotetica tecnologia, sarebbe (senza meno) doveroso interrogarsi sul senso profondo di questo tipo di intervento. Non che la questione non si ponga affatto nella prima ipotesi, ma in questa seconda è indubbiamente ancor più evidente. Nel caso allora della persona *nata* con grave disabilità mentale la *quaestio* è così articolata: l'ausilio tecnologico è funzionale a "guarire" questa persona o è volta a "trasformarla" in un'altra persona che non è e non è mai stata o, peggio, in *qualcos'altro* che non è, né è mai stata? Soprattutto, una persona con disabilità è da considerarsi una persona "malata da guarire"? O è una persona che si trova semplicemente in una delle (tante) «condizioni personali», cui fa riferimento l'articolo 3 della Costituzione? In questo caso l'impiego dell'intelligenza artificiale è veramente diretto al servizio delle vulnerabilità? Con il suo utilizzo si intende facilitare l'inclusione, la partecipazione al consesso sociale di questa persona con disabilità o s'intende uniformarla ad uno *standard* di umanità indispensabile per essere riconosciuta come persona umana?

Provo a spiegarmi meglio.

Com'è noto la Sindrome di Down o Trisomia 21 è una particolare condizione genetica caratterizzata dalla presenza di un terzo cromosoma nella coppia n. 21. Questo cromosoma "aggiuntivo" provoca, tra l'altro, una disabilità intellettiva variabile da persona a persona e dei tipici tratti somatici che in passato hanno indotto ad usare l'orrendo termine "mongoloide" per riferirsi appunto alle persone con Sindrome di Down; un termine purtroppo usato ancor'oggi in senso dispregiativo che nell'offendere queste persone offende altresì le valorose genti della Mongolia[43].

[43] A riguardo ed anche sugli stereotipi da superare relativi alle persone con Sindrome di Down vedi ora G. MATUCCI, *Parole per includere. Parole per educare. Rileggendo Cosimo Laneve*, in G. LANEVE (a cura di), *Parola, persona, comunità. Rileggendo* Parole per educare *di Cosimo Laneve*, eum - Edizioni Università di

Sono stati fatti studi diretti a "silenziare" il terzo cromosoma della coppia n. 21 (il cromosoma in più), onde tentare almeno di "arginare" le conseguenze "negative" che la sua presenza nel codice genetico determina[44].

La Sindrome di Down, però, pur con tutti i problemi che comporta, *in primis* a chi vive questa condizione, non è una malattia, ma è appunto una "condizione umana", una delle «condizioni personali» in relazione alle quali l'art. 3 Cost. sancisce il divieto di discriminazione[45].

Ammettiamo, per ipotesi, che sia effettivamente possibile azzerare gli effetti prodotti dalla presenza del cromosoma in più. Non possiamo sottrarci, allora, dal riflettere su quale sia il senso di un intervento di questo tipo. Si tratta di provare a "guarire" una persona con la Sindrome di Down? Si tratta di aiutare questo essere umano a stare oggettivamente meglio o lo stiamo "trasformando" in un'altra persona che non è? Soprattutto, (ancora una volta) se la Sindrome di Down non è una "malattia", chi vive questa condizione necessita di essere "guarito"?

Macerata, Macerata, 2025, 45 ss.

[44] Su tale tecnica vedi le considerazioni di B. DALLAPICCOLA: *Lo "spegnimento" di un cromosoma 21 nelle cellule della Sindrome di Down*, 2013 (*https://www.aipdroma.it/lo-spegnimento-di-un-cromosoma-21-nelle-cellule-della-sindrome-di-down/*). Oggi pare che almeno teoricamente sia addirittura possibile eliminare il cromosoma in più: *Trisomic rescue via allele-specific multiple chromosome cleavage using CRISPR-Cas9 in trisomy 21 cells*, in «PNAS Nexus», 2025, Vol. 4, n. 2 (18 February 2025). Sui dati genetici qualificati come «dati *sensibilissimi* destinatari di un elevato livello di protezione» vedi L. CHIEFFI, *Questioni di Biodiritto in una prospettiva europea*, in «Rassegna di diritto pubblico europeo», 2, 2022, 285.

[45] G. MATUCCI, *Parole per includere. Parole per educare. Rileggendo Cosimo Laneve*, cit., 43-44: «La disabilità altro non è che una mera condizione personale, quale può essere l'essere donna, omosessuale, o il trovarsi in una situazione di povertà economica: una condizione della quale l'ordinamento deve tenere conto per assicurare a chi vi si trovi il *pari* godimento dei diritti fondamentali».

Tutti gli interrogativi sopra formulati hanno una caratura bioetica e, per quel che qui più interessa, biogiuridica[46] così densa da far «tremar le vene e i polsi»[47].

Ed è proprio l'incessante accelerazione nelle conoscenze tecnico-scientifiche a far sorgere problemi di tale portata, ad imporci di raccogliere la «sfida […] di trovare nello strumentario del diritto soluzioni che fungano da mezzo di gestione delle complessità prodotte dagli avanzamenti tecnologici e che siano capaci di bilanciare le esigenze di tutela delle persone, l'innovazione, la cura e la ricerca, evitando che i diritti dei più deboli restino alla mercé di un determinismo tecnologico, che classificherebbe come lecito tutto ciò che è praticamente possibile»[48].

Qual è il discrimine tra, da un lato, *aiutare*, *sostenere*, *includere* e, dall'altro, *trasformare*, *assimilare*, *non accettare le diversità*?

Il biodiritto serve appunto a fissare positivamente un confine, a imporre cioè un limite al dominio della tecnica e il "cyber-biodiritto" nel contesto attuale, che tende sempre più ad adattarsi all'*agere*, proprio dell'intelligenza artificiale, e a marginalizzare l'*intelligere*, che qualifica autenticamente gli esseri umani[49], è chiamato, in particolare, a salvaguardare i confini dell' "umano",

[46] Per gli elementi di connessione e di distinzione tra bioetica e biodiritto vedi G. FONTANA, *Ricerca scientifica e libertà di cura. Scientismo ed antiscientismo nella prospettiva costituzionale*, Editoriale Scientifica, Napoli, 2019, 140 ss.

[47] Per dirla come il Sommo Poeta: *Inferno*, Canto I, verso 90 [D. ALIGHIERI, *La Divina Commedia. Inferno*, a cura di N. Sapegno, La Nuova Italia, Scandicci (Firenze), 1985 (3ª ristampa: gennaio 1987), 12].

[48] Così M. TOMASI, *Genetica e Costituzione. Esercizi di eguaglianza, solidarietà e responsabilità*, Editoriale Scientifica, Napoli, 2019, 371.

[49] Vedi L. FLORIDI, Agere sine intelligere. *L'intelligenza artificiale come nuova forma di agire e i suoi problemi etici*, in L. FLORIDI, F. CABITZA, *Intelligenza artificiale. L'uso delle nuove macchine*, Giunti Editore/Bompiani, Firenze-Milano, 2021, 113 ss. Vedi anche L. FLORIDI, *Etica dell'intelligenza artificiale. Sviluppi, opportunità sfide*, cit., *passim*.

dato che «le delicate questioni che si affacciano nell'era del post-umano […] già vengono all'attenzione del giurista per l'erompere di *robot* potenzialmente capaci di assumere decisioni autonome e di interagire con terzi»[50]. In fondo, la «grande sfida» delle scienze umane oggi consiste proprio nella «ricostruzione e rielaborazione dell'umano»[51].

5. L'umanità intera non è un "gruppo omogeneo"

Sciogliere questi nodi è certamente impresa titanica.

Per quanto concerne la posizione delle persone con disabilità un punto ragionevole d'avvio nell'approcciarsi a tali questioni potrebbe consistere nel far riferimento al canone ispiratore della Convenzione delle Nazioni Unite sui diritti delle persone con disabilità (ratificata e resa esecutiva in Italia con legge n. 18/2009), espresso nell'adagio «*nothing about us*, *without us*», che sancisce la doverosità del coinvolgimento dei diretti interessati nelle scelte che li riguardano e che vede una sua significativa declinazione nel paradigma dell'«autonomia protetta» approfondito e valorizzato in dottrina[52].

Del resto, il Regolamento europeo stesso promuove un modello regolatorio improntato ad un metodo di tipo partecipativo. Si pensi, ad esempio, agli articoli 56 (Codici di buone pratiche)[53], 57

[50] Vedi M. RUOTOLO, *Corso di diritto costituzionale*, 4ª Edizione, Giappichelli, Torino, 2026, 227.

[51] Vedi G. LANEVE, *Parole e diritto (costituzionale): alcune note introduttive*, in G. LANEVE (a cura di), *Parola, persona, comunità. Rileggendo* Parole per educare *di Cosimo Laneve*, eum - Edizioni Università di Macerata, Macerata, 2025, 29.

[52] Vedi G. MATUCCI, *Persona, formazione, libertà. L'autorealizzazione della persona con disabilità fra istruzione e* legal capacity, Franco Angeli, Milano, 2021.

[53] Art. 56, comma 3, Regolamento (UE) 2024/1689 del Parlamento Europeo

(Spazi di sperimentazione normativa per l'IA)[54], 62 (Misure per i fornitori e i deployer, in particolare le PMI, comprese le start-up)[55] e 67 (Forum consultivo)[56].

e del Consiglio: «L'ufficio per l'IA può invitare tutti i fornitori di modelli di IA per finalità generali, nonché le pertinenti autorità nazionali competenti, a partecipare all'elaborazione dei codici di buone pratiche. Le *organizzazioni della società civile*, l'industria, il mondo accademico e altri pertinenti portatori di interessi, quali i fornitori a valle e gli esperti indipendenti, possono sostenere il processo» (corsivo mio).

[54] Art. 57, comma 17, Regolamento (UE) 2024/1689 del Parlamento Europeo e del Consiglio: «La Commissione sviluppa un'interfaccia unica e dedicata contenente tutte le informazioni pertinenti relative agli spazi di sperimentazione normativa per l'IA per permettere ai portatori di interessi di interagire con gli spazi di sperimentazione normativa per l'IA e di formulare richieste di informazioni presso le autorità competenti e di chiedere orientamenti non vincolanti sulla conformità di prodotti, servizi e modelli aziendali innovativi che integrano le tecnologie di IA, a norma dell'articolo 62, paragrafo 1, lettera c). La Commissione si coordina proattivamente con le autorità nazionali competenti, se del caso». Sugli spazi di sperimentazione normativa vedi F. SANCHINI, *Le* sandboxes *normative tra evoluzione tecnologica e profili di compatibilità costituzionale*, in «AmbienteDiritto.it», 3, 2025; da ultimo, in raffronto alla «decretazione integrativa e correttiva», A. STEFANONI, *Decretazione integrativo-correttiva e* regulatory sandboxes *nel sistema delle fonti: profili vecchi e nuovi di una normazione a più fasi*, in «Consulta Online», II, 2026, 844 ss.

[55] Art. 62, comma 1, lett. d), Regolamento (UE) 2024/1689 del Parlamento Europeo e del Consiglio: «Gli Stati membri intraprendono le azioni seguenti: […] d) agevolare la partecipazione delle PMI e di altri portatori di interessi pertinenti al processo di sviluppo della normazione».

[56] Art. 67, commi 1 e 2, Regolamento (UE) 2024/1689 del Parlamento Europeo e del Consiglio: «1. È istituito un forum consultivo per fornire consulenza e competenze tecniche al consiglio per l'IA e alla Commissione, nonché per contribuire ai loro compiti a norma del presente regolamento. // 2. La composizione del forum consultivo rappresenta una selezione equilibrata di portatori di interessi, tra cui l'industria, le start-up, le PMI, la società civile e il mondo accademico. La composizione del forum consultivo è equilibrata per quanto riguarda gli interessi commerciali e non commerciali e, all'interno della categoria degli interessi commerciali, per quanto riguarda le PMI e le altre imprese».

Resta il fatto che, come messo in luce anche dalla stessa Corte costituzionale, le persone con disabilità «non costituiscono un gruppo omogeneo»[57]; e di questo è consapevole anche il legislatore eurounitario: basti qui richiamare ancora una volta il tenore testuale del Considerando n. 80 del Regolamento laddove prevede che occorre «tenere pienamente conto delle loro dignità e diversità intrinseche».

Il punto è che è l'umanità stessa a non essere un "gruppo omogeneo"; quindi anche le persone con disabilità partecipano di questa peculiare natura che connota l'intero *genere umano*, che in quanto *tale* ha appunto dentro di sé tante *species*. Con ciò, ovviamente, fermo che nemmeno «il concetto di "popolo"» va inteso «in un senso etnico»[58], non si intende certo far riferimento ad inesistenti, fantomatiche, differenti "razze". In proposito, basti por mente al noto aneddoto di einsteiniana memoria, per cui il grande scienziato nel 1933, al momento del suo ingresso negli Stati Uniti d'America, quale "razza" di appartenenza avrebbe indicato quella "umana", l'unica di cui avesse conoscenza. Semmai, si vuole sottolineare come il genere umano sia al suo interno ricco di diversità, che, come tali, vanno riconosciute e valorizzate, dato che «più noi ci rendiamo conto che le differenze sono importanti, per quel che richiamano in termini anche valoriali di rispetto e di importanza della persona, più la società in qualche modo diventa per così dire più umana»[59].

Le persone con disabilità, «la più grande minoranza che esiste

[57] Vedi Corte cost., 26 febbraio 2010, n. 80, n. 3 del *Considerato in diritto*.

[58] Cfr. J. HABERMAS, *L'inclusione dell'altro. Studi di teoria politica* [1996], a cura di L. Ceppa, Feltrinelli, Milano, 2013, 155.

[59] Vedi M. COLASANTO, *Conclusioni*, in F. GIRELLI (a cura di), *Lavoro e disabilità. Disciplina normativa e percorsi di inserimento*, Atti del Seminario, Roma, 12 giugno 2008, Editoriale Scientifica, Napoli, 2010, 113.

al mondo»[60], *a pieno titolo* (purtroppo non è pleonastico ribadirlo), così come sono, proprio in quanto esseri umani, appartengono a questo "gruppo *non* omogeneo" che è l'umanità; insomma, la realtà è che siamo tutti diversi e «abbiamo tutti diritto a conservare la nostra diversità, la nostra identità»[61]: è lo stesso principio di eguaglianza ad esigere «che a diversità di situazioni corrisponda diversità di trattamento»[62].

Ora, riuscire a far convivere le diversità è l'obiettivo di un percorso certamente non semplice, ma da intraprendere inesorabilmente[63]: le principali coordinate costituzionali in questo senso si possono rinvenire senz'altro nei doveri inderogabili di solidarietà sociale di cui all'art. 2 Cost. e nel divieto di discriminazioni in ragione di condizioni personali e sociali, sancito nell'art. 3 Cost., che, come noto, impone altresì la rimozione degli ostacoli che «di fatto»

[60] Così G. DE SANTIS, *Lavoro e disabilità: una nuova linea di riflessione e ricerca*, in F. GIRELLI (a cura di), *Lavoro e disabilità. Disciplina normativa e percorsi di inserimento*, Atti del Seminario, Roma, 12 giugno 2008, Editoriale Scientifica, Napoli, 2010, 20.

[61] Vedi G. M. FLICK, *Relazione conclusiva. Diritto e disabilità o "diritto alla disabilità"?*, in C. COLAPIETRO e A. SALVIA (a cura di), *Assistenza, inclusione sociale e diritti delle persone con disabilità. A vent'anni dalla legge 5 febbraio 1992, n. 104*, Editoriale Scientifica, Napoli, 2013, 391. Ad avviso di V. PAGNANELLI, *Non discriminazione e diritto alla diversità: cosa c'è di nuovo per i disabili nell'era dell'AI?*, in «BioLaw Journal – Rivista di BioDiritto», Special Issue 1, 2024, 344 ss. un corretto impiego dell'intelligenza artificiale, nella cornice della disciplina eurounitaria, consoliderebbe la tutela del «diritto alla diversità» delle persone con disabilità.

[62] Vedi C. MORTATI, *Istituzioni di diritto pubblico*, II, Settima edizione rielaborata ed aggiornata, Cedam, Padova, 1967, 828.

[63] Ed un significativo contributo in concreto auspicabilmente potrà esser fornito in questa direzione dal neoistituito Garante nazionale dei diritti delle persone con disabilità, su cui, se si vuole, F. GIRELLI, *L'Autorità «Garante nazionale dei diritti delle persone con disabilità»*, in E. VIVALDI (a cura di), *Il progetto di vita della persona con disabilità. Dal PNRR al decreto legislativo n. 62/2024*, Pisa University Press, Pisa, 2025, 85 ss.

limitano la libertà e l'eguaglianza dei cittadini: in fondo, l'«aspirazione dell'eguaglianza sostanziale è proprio quella di favorire» chi versa in una qualche condizione di vulnerabilità, «che di fatto ne impedisce il pieno sviluppo della personalità e l'effettiva partecipazione alla vita comunitaria»[64].

Del resto, il programma costituzionale repubblicano vede la realizzazione della «pari dignità sociale» come funzionale al «pieno sviluppo della persona umana»[65], atteso che la prima implica che lo Stato si attrezzi per «operare contro la miseria e contro la ignoranza e contro il delitto e contro gli altri mali che degradano la dignità sociale dei cittadini»[66].

Se, allora, il principio personalista innerva l'intera trama del tessuto costituzionale, non c'è ragione *costituzionale*, appunto, per cui da tutto ciò *alcune persone* vengano escluse, soprattutto nell'attuale contesto, ove si dà (e sempre più si darà) applicazione ad una tecnologia così pervasiva, quale è l'intelligenza artificiale, che si mostra anzi così potenzialmente incisiva su quello che ad oggi intendiamo ancora per "umanità".

La condizione di vulnerabilità è propria dell'intero genere umano sin dall'inizio dei tempi, «è un elemento costitutivo della condizione antropica». Della vulnerabilità, quindi, giova precisarlo, non si può che avere «una concezione relativa», poiché dipende dai soggetti messi a raffronto e non necessariamente costituisce una condizione permanente[67].

In fondo la «nozione di vulnerabilità si presta a svariate letture,

[64] Cfr. Q. CAMERLENGO, *Eguaglianza*, in S. ILLARI (a cura di), *Eguaglianza. Una sfida per le società plurali*, Carocci, Roma, 2025, 51.

[65] F. MODUGNO, *I diritti del consumatore: una nuova «generazione» di diritti?*, in *Scritti in onore di Michele Scudiero*, III, Jovene, Napoli, 2008, 1371.

[66] Vedi C. ESPOSITO, *Eguaglianza e giustizia nell'art. 3 della Costituzione*, in ID., *La Costituzione italiana. Saggi*, Cedam, Padova, 1954, 62.

[67] Vedi M. LUCIANI, *Le persone vulnerabili e la Costituzione*, cit.

tutte incentrate sulla persona umana, colta nelle diverse fasi della vita e nelle relazioni con gli altri»[68]. La stessa Corte europea dei diritti dell'uomo ha dato una «connotazione volutamente aperta del concetto di vulnerabilità», cui possono essere ricondotte anche le persone con disabilità[69], che, anzi, ne rappresentano una figura emblematica, come ha chiarito a suo tempo lo stesso presidente della Corte europea dei diritti dell'uomo[70].

Le cosiddette "situazioni di vulnerabilità", pertanto, sono quelle proprie di persone *più* vulnerabili o, se si vuole, più deboli, meno forti di altre[71], la cui vita naturalmente è connotata nel profon-

[68] Così S. SCIARRA, *Intervento* svolto in occasione della *lectio magistralis* del Presidente della Corte europea dei diritti dell'uomo tenuta a Roma all'Università "La Sapienza" il 22 aprile 2022 (*https://www.cortecostituzionale.it/documenti/convegni_seminari/sciarra_discussant_spano_finale_20220503170804.pdf*).

[69] Vedi P. SCARLATTI, *Tutela dei diritti e trattamento dei detenuti vulnerabili. A proposito del recente caso* Sy contro Italia, in «Dirittifondamentali.it», 1, 2022, 535-537.

[70] Vedi R. SPANO, *Diritti umani e persone vulnerabili*, *Lectio magistralis* tenuta a Roma all'Università "La Sapienza" il 22 aprile 2022 (*https://www.cortecostituzionale.it/documenti/convegni_seminari/itervento_spano_20220503170732.pdf*).

[71] Sul concetto di soggetto «debole» quale categoria giuridica vedi: M. AINIS, *I soggetti deboli nella giurisprudenza costituzionale*, in «Politica del diritto», 1, 1999, 25 ss.; L. AZZENA, *Divieto di discriminazione e posizione dei soggetti «deboli». Spunti per una teoria della «debolezza»*, in C. CALVIERI (a cura di), *Divieto di discriminazione e giurisprudenza costituzionale*, Atti del Seminario di Perugia del 18 marzo 2005, Giappichelli, Torino, 2006, 35 ss. Per la peculiare prospettiva della vulnerabilità vedi P. SCARLATTI, *I diritti delle persone vulnerabili*, Editoriale Scientifica, Napoli, 2022; ID., *Soggetti deboli, Costituzione ed istanze della vulnerabilità*, in «Gruppo di Pisa. Dibattito aperto sul Diritto e la Giustizia Costituzionale», 1, 2023, 266 ss.; C. SEVERINO, H. ALCARAZ (dir.), *Systèmes de contrôle de constitutionnalité par voie incidente et protection des personnes en situation de vulnérabilité. Approche de droit comparé*, DICE Éditions, Aix-en-Provence, 2021 (Date de mise en ligne: 2 juin 2022). Per la distinzione tra «vulnerabilità universale» e «vulnerabilità particolare» vedi S. DADÀ, *Vulnerabilità. Note sul ruolo del*

do dalla dignità, che, così come per qualsiasi altro essere umano, «mai può essere perduta»[72]. E proprio perché ancor'oggi fatica ad esaurirsi l'approccio culturale per cui da una parte ci sono i cosiddetti "normali" e dall'altra i "disabili", stenta a consolidarsi la (generalizzata) consapevolezza del fatto che la disabilità è una delle tante manifestazioni della persona umana, collocata al centro del programma costituzionale repubblicano.

La chiave fondamentale di tutela giuridica delle persone con disabilità, appunto perché *persone* e senza retorica alcuna (in tema di disabilità, del resto, serve soprattutto senso pratico), non può che rinvenirsi, allora, nella salvaguardia della dignità umana, «punto di equilibrio» tra eguaglianza e libertà[73] e canone interpretativo di difesa dalle insidie della tecnica[74], atteso che «le vulnerabilità impediscono alle persone di rialzare la testa, ma la dignità dell'uomo impedisce di chinarla»[75].

concetto nell'AI Act, in «BioLaw Journal – Rivista di BioDiritto», Special Issue 1, 2024, 41 ss., la quale ritiene che il Regolamento europeo recepisca la concezione della vulnerabilità «principalmente in senso particolare» (52).

[72] Vedi M. RUOTOLO, *Corso di diritto costituzionale*, cit., 228.

[73] Sul punto vedi A. PIROZZOLI, *La dignità dell'uomo. Geometrie costituzionali*, Edizioni Scientifiche Italiane, Napoli, 2012, 24-25.

[74] Vedi L. CHIEFFI, *Dignità umana e sviluppi del principio personalista. Brevi note introduttive*, in L. CHIEFFI (a cura di), *Dialoghi sulla dignità umana*, in «Rassegna di diritto pubblico europeo», 1, 2013, 6 ss. Vedi anche A. SIMONCINI, *L'algoritmo incostituzionale: intelligenza artificiale e il futuro delle libertà*, cit., 88: «L'unica strada possibile, dunque, è far sì che il valore (dunque il senso) della tutela della dignità della persona e della sua libertà, divengano parte integrante della formazione di coloro che poi elaboreranno quelle tecnologie».

[75] Vedi A. GENTILI, *La vulnerabilità sociale. Un modello teorico per il trattamento legale*, in «Rivista Critica del Diritto Privato», 1, 2019, 64.

Paradigmi organizzativi e di apprendimento in Agenzia delle Entrate. L'impatto della digitalizzazione

di *Salvatore Lazzara*[*]

SOMMARIO: 1. Modelli post-burocratici per la fiscalità – 2. Verso un'idea di valore pubblico – 3. Apprendimento organizzativo e nuove tecnologie – 4. Conclusioni.

1. Modelli post-burocratici per la fiscalità.

Il processo di riforma amministrativa negli ultimi trent'anni ha visto l'affacciarsi di alcuni paradigmi post-burocratici[1] che sono riusciti ad imporsi quale alternativa ai tradizionali sistemi di *government* (Di Mascio e Natalini, 2022)[2]. In queste pagine partiremo dall'analisi dell'affermazione ed evoluzione, nel contesto italiano, del *New public management*, rispetto al quale metteremo al centro il processo di *Agencification* e gli esiti che quest'ultimo ha avuto nel sistema della fiscalità. Riteniamo che, al netto di un odierno indebolimento della pervasività e capacità esplicativa del paradigma NPM, delle affermazioni in ambito accademico sulla

[*] Dottorando in Società in mutamento: politiche, diritti e sicurezza, Università degli Studi della Tuscia.

[1] Il riferimento al termine *paradigma* è preso in prestito dal lavoro di THOMAS KUHN, *La struttura delle rivoluzioni scientifiche* (trad. it., Einaudi, Torino, 2009); è qui inteso come il complesso di regole metodologiche, modelli esplicativi, criteri di soluzione dei problemi, predominante in un determinato periodo storico e riconosciuto come efficace da una comunità.

[2] Fra i più rilevanti: il *new public management*, il *neo weberian state*, la *new public governance*, *l'open government*, la *digital era governance*.

sua crisi e decadenza (Dunleavy *et alia*, 2006; Borioni, 2017; Moini, 2017), nel contesto istituzionale italiano e in particolare nel settore della fiscalità il paradigma mostri non solo una forte *resilienza* ma una capacità di costante innovazione amministrativa; tale vivacità si giustifica in ragione di costanti processi di ibridazione, tanto con altri paradigmi concorrenti, quanto con le culture preesistenti (Pommier Vincelli, 2018).

Da un lato, infatti, l'orientamento al *New public management* tipico delle agenzie fiscali (Ongaro, 2006) continua a rappresentare una sorta di precondizione strategica all'azione di queste nell'ambito della propria missione istituzionale (Franceschetti, 2022), soprattutto attraverso la permanenza, nel modello organizzativo di riferimento, degli elementi strutturali essenziali delle *Next steps agencies* (Arena, 1998). Dall'altro, la traiettoria evolutiva delle agenzie fiscali si connota per aver intercettato altri paradigmi di riforma, in una tensione verso l'ibridazione che ha dato modo a quel modello organizzativo di mutare gli elementi meno essenziali o non funzionali, mantenendo coerenza ed efficacia[3].

Le Agenzie fiscali sono state create per svolgere un'attività caratterizzata da un alto grado di misurabilità dei risultati sia in termini di *output* (qualità dei servizi erogati), che in termini di *outcome*, cioè di impatto della propria azione sulla collettività, che vanno intesi come azioni per la riduzione dell'evasione fiscale e l'incremento costante della *tax compliance*. Per essere compresa appieno, l'esperienza delle agenzie fiscali va considerata però nell'ambito del più generale interesse delle Pa per i temi del management, mutuati dal settore privato, sviluppatosi nel corso di una lunga stagione di riforme amministrative avviata dalle cosiddette

[3] In queste pagine analizzeremo più nel dettaglio l'esperienza dell'Agenzia delle Entrate, non potendo procedere ad una comparazione puntuale dell'evoluzione delle diverse Agenzie.

riforme Bassanini. Elementi distintivi di questo orientamento sono rintracciabili nei meccanismi di relazione *Principal-agent* creati *ad hoc* per le agenzie fiscali, come la Convenzione con il Ministero dell'Economia e delle Finanze e la creazione di sistemi di *budgeting e reporting* volti al controllo dei processi operativi, a misurare e valutare la performance organizzativa e del personale ai vari livelli e lo sviluppo di sistemi professionali per l'evoluzione e la manutenzione delle competenze.

Nella fase di prima istituzione le agenzie fiscali hanno ricevuto dal D. Lgs. n. 300/99 una disciplina *ad hoc*[4]. Per esse il decreto ha definito un sistema di autonomie più accentuato rispetto a quanto stabilito dagli artt. 8 e 9 per le agenzie in generale. In particolare, l'art. 61 del decreto riconosce loro esplicitamente autonomia gestionale, regolamentare, amministrativa, patrimoniale, organizzativa, contabile, finanziaria e tecnica, assoggettando le medesime al controllo di competenza della Corte dei conti. Oltre alla disciplina generale dettata dagli artt. 8 e 9 vengono compiute specifiche previsioni in merito alla natura e al contenuto degli statuti (art. 66), agli organi interni e alle funzioni a essi attribuite (art. 67), alla disciplina contabile e finanziaria (art. 70) e alla gestione delle fasi del cambiamento (art. 73). Un ulteriore tratto distintivo delle agenzie fiscali riguarda l'attribuzione di una propria personalità giuridica di diritto pubblico. Per quanto riguarda l'inquadramento contrattuale del personale, per le agenzie fiscali era stato inizialmente istituito un comparto autonomo, confluito però nel CCNL Funzioni centrali a partire dalla tornata contrattuale 2016 – 2018.

[4] Il D.lgs n. 300/99 imprime al processo di *agencification* una fortissima accelerazione. Il decreto si proponeva infatti lo scopo di promuovere una riforma complessiva dell'organizzazione del governo. Fra le principali novità proposte vi era l'istituzione di undici nuove agenzie, la maggior parte delle quali non ha tuttavia mai visto la luce, anche a causa del mutato orientamento dei governi successivi.

Va detto che le quattro (ora tre) agenzie fiscali[5] istituite dal Decreto, pur condividendo missioni istituzionali afferenti al medesimo macrocontesto[6], presentano dal punto di vista organizzativo un modello non del tutto uniforme; le caratteristiche dei diversi settori di *policy* e le singole vicende amministrative hanno fatto sì che si differenziassero, almeno in parte[7]. Anche questa variabilità di schema organizzativo fa delle agenzie fiscali una forma *resiliente e ibridata* del modello *idealtipico* per agenzie previsto dal paradigma del NPM, nel confronto con le applicazioni *pure* del modello *Next Steps* nato nel Regno Unito e rispetto ad altre esperienze simili, nate nello stesso periodo e a partire dalla medesima previsione normativa.

Uno sguardo retrospettivo ci porta ad affermare che, a venticinque anni dalla sua affermazione, il processo di *agencification* può in larga parte considerarsi il frutto di una stagione di innovazione organizzativa conclusa. Tuttavia, riteniamo che il modello di agenzia pubblica, applicato all'amministrazione della fiscalità, riveste ancora oggi una particolare importanza, alla luce della sua capacità di far coesistere elementi tipici del NPM in accostamento ad altri paradigmi (Pollitt e Bouckaert, 2011) quali la *Digital Era Go-*

[5] Agenzia delle Entrate, Agenzia del Demanio, Agenzia delle Dogane e dei Monopoli. L'Agenzia del Territorio, anch'essa prevista dal Decreto n. 300/99, con il Decreto-legge del 6 Luglio 2012, è stata incorporata dall'Agenzia delle Entrate. Il 1° dicembre 2012 l'Amministrazione Autonoma dei Monopoli di Stato (AAMS) è stata fusa con l'Agenzia delle Dogane, creando l'attuale Agenzia delle Dogane e dei Monopoli (ADM), come stabilito dal D.L. n. 87 del 27 giugno 2012, convertito con la Legge 7 agosto 2012, n. 135.

[6] Alle agenzie, infatti, compete l'esercizio di compiti tecnico-operativi che prima erano esercitati dal ministero di riferimento.

[7] È il caso dell'agenzia del Demanio, che per effetto del D.Lgs. n. 173 del 2003 si trasforma da amministrazione dotata di personalità giuridica di diritto pubblico a Ente pubblico economico, intervenendo in tal senso a modifica dell'art. 61 del D.Lgs. n. 300/1999.

vernance, l'Open government e il *Neo Weberian State.* Sotto quest'ultimo profilo, l'ibridazione più significativa forse consiste nell'avere incorporato, all'interno dello schema tipico del NPM, elementi del paradigma del *Valore pubblico* (Moore, 2005).

2. Verso un'idea di valore pubblico

In linea generale il percorso italiano verso il paradigma del VP trova la sua prima sistematizzazione nelle *Linee Guida per il Piano della Performance*[8] (Deidda Gagliardo, 2021). Esse rappresentano un tentativo di superamento della logica puramente quantitativa della misurazione, prevista dal ciclo della performance introdotto dalla riforma Brunetta (D.lgs. n.150/2009). Il riferimento alle Linee guida è qui fondamentale e non solo per questioni definitorie. Se infatti queste definiscono il Valore Pubblico come il *miglioramento del livello di benessere economico-sociale rispetto alle condizioni di partenza della politica o del servizio*, il concetto stesso di Valore Pubblico viene messo in stretta relazione con il ciclo della performance, specificando che la Pubblica Amministrazione realizza Valore Pubblico quando persegue *un miglioramento congiunto ed equilibrato degli impatti esterni ed interni delle diverse categorie di utenti e stakeholder, favorendo la possibilità di generarlo anche a favore di quelli futuri.*

Questa definizione evidenzia l'adozione di un approccio *multidimensionale* che integra efficacia, efficienza ed economicità con

[8] Elaborate dal Dipartimento della Funzione Pubblica a partire dal 2017 e redatte ai sensi dell'articolo 3, comma 1, del D.P.R. 105/2016 che attribuisce al DFP le funzioni di indirizzo, coordinamento e monitoraggio in materia di ciclo della performance, avvalendosi del supporto tecnico e metodologico della Commissione Tecnica per la Performance (CTP) di cui all'articolo 4 del citato decreto.

dimensioni più ampie di impatto sociale (Cepiku e Meneguzzo, 2011). La generazione del benessere complessivo e multidimensionale parte quindi dal definire obiettivi di Valore Pubblico precisi (nel contenuto, nel tempo e nello spazio) e mettere in relazione con questi gli obiettivi di performance (che agiscono come leve di creazione del VP), le misure di contenimento dei rischi – non solo corruttivi – (che agiscono come leve di protezione del VP), le azioni di miglioramento della salute delle risorse (che agiscono come condizioni abilitanti del VP).

Con il Decreto-legge n. 80 del 9 giugno 2021[9], recante *Misure urgenti per il rafforzamento della capacità amministrativa delle pubbliche amministrazioni funzionale all'attuazione del Piano nazionale di ripresa e resilienza (PNRR) e per l'efficienza della giustizia*, il paradigma del VP diviene infine centrale nel processo di innovazione amministrativa, attraverso l'introduzione del Piano Integrato di Attività e di Organizzazione (PIAO), strumento finalizzato a integrare gli strumenti di programmazione orientandoli alla creazione di Valore Pubblico (Deidda Gagliardo, 2021, cit.)[10]. Nel PIAO confluiscono una serie di documenti/artefatti che avevano una propria autonomia con riferimento alle tempistiche, ai contenuti e alle norme di riferimento: il Piano dei Fabbisogni di Personale (PFP), il Piano delle azioni concrete (PAC), il Piano per Razionalizzare

[9] Convertito dalla legge n. 113/2021.

[10] L'iter legislativo riguardante il PIAO è stato definitivamente completato con la pubblicazione sulla Gazzetta Ufficiale n. 151 del 30 giugno 2022, del D.P.R. n. 81 "Regolamento recante individuazione degli adempimenti relativi ai Piani assorbiti dal Piano integrato di attività e organizzazione" e con la sottoscrizione – da parte del Ministro per la Pubblica Amministrazione, di concerto con il Ministro dell'Economia e delle Finanze – del decreto ministeriale del 30 giugno 2022 n. 132, che definisce i contenuti e lo schema tipo del Piano, nonché le modalità semplificate per la sua adozione da parte degli enti con meno di 50 dipendenti.

l'utilizzo delle Dotazioni Strumentali (PRSD), il Piano della Performance (PdP), il Piano triennale di Prevenzione della Corruzione e della Trasparenza (PtPCT), il Piano Organizzativo del Lavoro Agile (POLA) e il Piano di Azioni Positive (PAP).

Il PIAO sembrerebbe allora rappresentare, data la sua impostazione strategica e la natura di *artefatto complesso*, un punto di confluenza per le influenze derivanti da diversi paradigmi di riforma amministrativa.

Possiamo rintracciare evidenze a supporto di questa tesi proprio a partire dalla lettura del PIAO dell'Agenzia delle Entrate. Nell'ultimo documento pubblicato[11] al momento in cui scriviamo, il concetto di valore pubblico è chiaramente richiamato e definito come *l'impatto generato dalle politiche dell'ente sul livello di benessere complessivo e multidimensionale* (economico, sociale, ambientale, sanitario, ecc.) *di cittadini e imprese, ottenuto governando le performance in tale direzione, a partire dalla cura della salute organizzativa e delle risorse dell'ente*. La creazione di valore pubblico esterno dipende dal miglioramento delle performance; la creazione di valore pubblico interno è, dunque, il presupposto per la creazione di valore pubblico esterno. Ora, partendo dal presupposto che il compito principale dell'Agenzia delle Entrate è il perseguimento del massimo livello di *compliance* fiscale, l'Agenzia delle Entrate sembra orientarsi, al fine di generare e monitorare il valore pubblico prodotto, secondo due direttrici: una interna, perseguendo i paradigmi di efficienza, efficacia ed economicità nelle scelte aziendali; l'altra esterna, valutando gli impatti sui portatori di interesse di riferimento. Il valore pubblico si genera quando viene garantito l'equilibrio ottimale delle due dimensioni.

[11] L'Agenzia - Piano integrato di attività e organizzazione - Agenzia delle Entrate

3. Apprendimento organizzativo e nuove tecnologie

Ma le ragioni della resilienza del modello delle Agenzie, nell'esperienza dell'Agenzia delle Entrate, sono da rintracciare esclusivamente nelle capacità di incarnare una visione coerente di Valore pubblico? A nostro avviso andrebbe indagata anche la capacità dell'organizzazione di generare apprendimento e conoscenza, al suo interno e per il sistema della fiscalità, e in questo senso di incarnare le logiche della *Learning organization*. Le manifestazioni di questa capacità sono tanto più evidenti quanto più si analizza la reazione dell'Agenzia delle Entrate ai cambiamenti intervenuti, sotto le spinte di tipo normativo e tecnologico, nelle logiche di produzione ed erogazione dei servizi e nei sistemi professionali e di sviluppo delle competenze del personale. Sono tutti ambiti in cui si registra una costante spinta alla digitalizzazione e l'ingresso forte dell'uso degli algoritmi di IA a supporto dei processi, fattori questi che hanno suscitato speranze ma anche perplessità (Lalli, 2022, Palazzo, 2024).

Nel contesto dell'Agenzia delle Entrate si tratta di leggere il ricorso alla tecnologia di IA come uno strumento di modernizzazione del sistema amministrativo e di semplificazione del funzionamento (Torchia, 2023), come strumento facilitatore dell'innovazione; ma anche di comprendere come governare la capacità degli algoritmi di generare *un ordine nei sistemi di conoscenze* che sottostanno alle decisioni (Sterpa, 2024) e di intuire le linee di evoluzione della stessa *intelligenza amministrativa*, vale a dire il sapere condiviso dall'organizzazione verso l'esterno e nell'organizzazione verso le persone al suo interno (Sgueo, 2024).

Partendo da questi assunti, è interessante comprendere le modalità con cui l'Agenzia delle Entrate sta affrontando le sfide della digitalizzazione e dell'impiego di algoritmi di IA nel contesto della

riforma fiscale. Dal percorso di innovazione condotto in Agenzia delle Entrate sul piano organizzativo e su quello dei nuovi servizi al cittadino/contribuente può derivare la capacità di sviluppare conoscenza individuale e collettiva, a partire dalle qualità emergenti del sistema, e di radicare pratiche efficaci di apprendimento organizzativo. L'evoluzione normativa, organizzativa e tecnologica che ha interessato l'Agenzia delle Entrate ha infatti avuto come presupposto anche l'esigenza di generare conoscenza organizzativa, apprendimento e sviluppo di competenze.

Ci riferiamo qui alle applicazioni nel campo dell'analisi del rischio fiscale e del supporto alle interazioni con gli utenti; strategie di innovazione che si muovono nell'ambito della IA cognitiva e di quella "di facilitazione" (Buono, 2022).

In generale l'analisi del rischio fiscale è un'attività che consente di incrementare il livello di conoscenza dell'ambiente esterno, tramite un uso efficiente dei dati a disposizione, così da ridurre il livello di asimmetria informativa in cui si trova ad operare l'Amministrazione finanziaria[12]. In questo contesto le moderne tecniche di *machine learning* e di intelligenza artificiale possono essere integrate con metodologie classiche di tipo statistico-econometrico, in modo da coniugare l'esigenza di flessibilità della modellazione dei dati con quella dell'interpretabilità dei risultati.

Sul fronte del supporto alle interazioni efficaci con l'utenza, le azioni dell'Agenzia delle Entrate si orientano invece allo sviluppo di *chatbot* per l'assistenza agli utenti, al *text mining* per la categorizzazione automatica o semiautomatica dei documenti, alla sperimentazione di algoritmi per facilitare la gestione e l'analisi, da par-

[12] Ministero dell'Economia e delle Finanze, *Relazione sull'economia non osservata e sull'evasione fiscale e contributiva*, anno 2024 (art. 10-bis.1 c. 3 legge 31 dicembre 2009, n. 196).

te dei funzionari, della documentazione pervenuta, efficientando i tempi di lavorazione.

4. Conclusioni

Abbiamo brevemente delineato un quadro in cui la chiave di successo consiste, in primo luogo, nell'integrazione di queste funzionalità all'interno dei processi e nella gestione della multidisciplinarietà della conoscenza prodotta dalla compresenza di sistemi di competenze provenienti da mondi diversi, fino ad ora non integrati in un contesto di PA. Sembra allora naturale chiedersi quali siano le prospettive di ulteriore evoluzione del modello organizzativo dell'Agenzia delle Entrate, se sarà ancora il frutto della spinta all'ibridazione fra diversi paradigmi di riforma postburocratici, e se saprà mantenersi in equilibrio tra orientamento agli obiettivi ed allo sviluppo dell'*accountability*, continuando a sviluppare pratiche di apprendimento efficaci e rispondendo in maniera coerente agli impatti della digitalizzazione. In questo contesto problematico, la pervasività degli algoritmi nelle pratiche gestionali e amministrative pone al sistema professionale ed alle competenze del personale dell'AdE significative sfide ed opportunità.

Indirizzi strategici e principi normativi sull'utilizzo dell'intelligenza artificiale nella Pubblica Amministrazione

di *Paolo Macri*[*]

SOMMARIO: 1. Evoluzione tecnologica, contesti applicativi e indirizzi strategici dell'Unione europea in materia di intelligenza artificiale. – 2. Dall'ecosistema di eccellenza al principio di fiducia: il quadro europeo verso l'AI Act. – 3. Il principio antropocentrico nella Legge n. 132 del 2025 sull'intelligenza artificiale. – 4. Le Linee guida AGID per l'adozione dell'intelligenza artificiale nella Pubblica Amministrazione.

1. Evoluzione tecnologica, contesti applicativi e indirizzi strategici dell'Unione Europea in tema di Intelligenza artificiale

Nel corso del biennio 2024-2025 la diffusione massiva dei sistemi di intelligenza artificiale (IA) nel mondo del lavoro ha messo in luce gli effetti trasformativi di queste soluzioni tecnologiche sia per gli ambiti organizzativi che per quelli decisionali, con significative ricadute sulla produttività e sull'efficienza dei processi. L'applicazione di sempre più sofisticati algoritmi di apprendimento (*Large Language Models*) a grandi quantità di dati ha consentito la piena operatività di sistemi di Intelligenza artificiale generativa, in grado di simulare funzioni cognitive umane come la pianificazione, l'apprendimento, il ragionamento, la creatività e di formula-

[*] Dottorando in Società in mutamento: politiche, diritti e sicurezza, Università degli Studi della Tuscia.

re analisi e proporre soluzioni a questioni di notevole complessità in assenza di un intervento umano diretto[1].

Questa accelerazione tecnologica, dirompente e apparentemente inarrestabile, costituisce solo la più recente evoluzione dei sistemi di apprendimento basati sulle "reti neurali" che, a partire dalla metà degli anni Cinquanta del secolo scorso, si sono lentamente evoluti nei prodotti commerciali oggi disponibili e che, nel tempo, hanno portato ricercatori e studiosi a contribuire alle iniziative regolatorie del Legislatore eurounitario e dei Legislatori nazionali.

In questo contesto, l'Unione Europea ha formulato indirizzi strategici di rilevante portata, che costituiscono l'esito di un processo pluriennale, in cui la Commissione, con l'adozione del *Libro bianco sull'intelligenza artificiale* del 2020[2], ha individuato nel settore pubblico il comparto "apripista per l'uso dell'IA", stabilendo con il *Piano Coordinato sull'intelligenza artificiale* del 21 aprile 2021[3] un approccio indirizzato all'eccellenza e alla fiducia.

I principi e gli indirizzi elaborati in sede eurounitaria appaiono particolarmente rilevanti, in quanto la Commissione, pur impegnandosi a favorire i progressi scientifici e a preservare la leadership tecnologica dell'Unione Europea nel settore, ha posto le basi

[1] Per una riflessione complessiva sull'introduzione dell'intelligenza artificiale nelle amministrazioni pubbliche *cfr.* B. MARCHETTI, *Intelligenza artificiale e pubblica amministrazione: l'importanza dei distinguo per comprendere rischi e opportunità*, in E. HAPPACHER, C. ZWILLING, P. BUSSJÄGER, W. OBWEXER, S. PAROLARI, (a cura di) (2025), *Digitalisierung und regionale Autonomie - Digitalizzazione e autonomia regionale*, Nomos, Baden Baden, 91–104.

[2] *Libro bianco sull'intelligenza artificiale. Un approccio europeo all'eccellenza e alla fiducia*, COM (2020) n. 65 del 19 febbraio 2020.

[3] Il *Piano coordinato sull'intelligenza artificiale - cfr.* COM (2021) n. 205 del 21 aprile 2021 – si basa sulla collaborazione instaurata tra la Commissione e gli Stati membri durante il piano coordinato del 2018 *cfr.* COM (2018) n. 758 del 7 dicembre 2018.

strategiche dello sviluppo e dell'impiego dei sistemi di IA nella prospettiva del miglioramento complessivo della vita dei cittadini europei, senza cedere a compromessi sul rispetto dei loro diritti.

Nella scelta delle opzioni strategiche per il migliore impiego dei sistemi di intelligenza artificiale, salvaguardando i diritti dei cittadini, la Commissione ha coinvolto tutte le parti interessate: dagli Stati membri alle istituzioni europee, dagli *stakeholder* (industrie, parti sociali, rappresentanti della società civile) ai cittadini europei, per consentire a questi ultimi di usufruire di una migliore fruizione dei servizi pubblici e agli *stakeholder* di avvalersi di nuove generazioni di prodotti e servizi.

Nel definire questo approccio, la Commissione ha tenuto presente la necessità di non frammentare il Mercato unico e, proprio per questo, ha stabilito un primo quadro di riferimento per allineare gli sforzi europei, nazionali e regionali, raccogliendoli in un "sistema di eccellenza", al fine di favorire la migliore interazione tra il pubblico e il privato nell'adozione dei sistemi di IA. Uno degli scopi è quello di realizzare un "ecosistema di fiducia" volto non solo a consentire ai cittadini di fruire delle applicazioni di IA ma anche a garantire alle imprese e alle organizzazioni pubbliche la certezza del diritto, necessaria per innovare sistemi e processi, utilizzando l'IA.

Tra le azioni previste, la Commissione ha ritenuto essenziale che fosse il settore pubblico, nel perseguimento delle attività istituzionali, a utilizzare prodotti e servizi basati sull'IA, individuando i settori dell'assistenza sanitaria e dei trasporti come quelli tecnologicamente maturi per un utilizzo dell'IA su larga scala, coinvolgendo gli altri operatori del servizio pubblico nell'utilizzo di sistemi di IA in materia di appalti[4].

[4] Cfr. *Libro bianco sull'intelligenza artificiale*, cit., p. 8: «(…) *È essenziale che le amministrazioni pubbliche, gli ospedali, i servizi di pubblica utilità e di trasporto,*

Questo indirizzo, volto a sviluppare un ecosistema di IA su scala europea, ha ricevuto ulteriore impulso nel 2022 con il Programma strategico per il decennio digitale 2030[5].

Il Programma, sviluppato per rafforzare la sovranità digitale europea, ha definito quattro obiettivi fondamentali da raggiungere entro il 2030: il primo rivolto a far crescere le competenze digitali della popolazione (per cui almeno l'80% della popolazione tra i 16 e i 74 anni dovrà raggiungere competenze digitali di base e il numero di specialisti in "Tecnologie dell'Informazione e della Comunicazione" dovrà raggiungere i 20 milioni di impiegati nel territorio dell'Unione), il secondo incentrato su infrastrutture digitali sicure e sostenibili, il terzo rivolto alla trasformazione digitale delle imprese (attraverso l'utilizzo di almeno una soluzione tra servizi di cloud computing, di gestione di Big Data e di IA) e, infine, il quarto rivolto alla digitalizzazione dei servizi pubblici[6].

Con riferimento a quest'ultimo obiettivo, il Programma strategico per il decennio digitale 2030 ha stabilito che, entro la scadenza stabilita, il 100% dei servizi pubblici principali dovrà essere accessibile online (consentendo alle imprese e ai cittadini all'interno

le autorità di vigilanza finanziaria e altri settori di interesse pubblico inizino rapidamente a utilizzare nelle loro attività prodotti e servizi che si basano sull'IA. Un'attenzione particolare sarà rivolta ai settori dell'assistenza sanitaria e dei trasporti, in cui la tecnologia è abbastanza matura da consentire una diffusione su vasta scala. (…) *la Commissione avvierà dialoghi settoriali aperti e trasparenti dando priorità agli operatori del servizio pubblico, delle amministrazioni rurali e dell'assistenza sanitaria, al fine di presentare un piano d'azione che faciliti lo sviluppo, la sperimentazione e l'adozione dell'IA. I dialoghi settoriali saranno utilizzati per elaborare uno specifico "Programma di adozione dell'IA", che sosterrà gli appalti pubblici di sistemi di IA e contribuirà a trasformare le procedure stesse degli appalti pubblici* (…)».

[5] Il *Programma* è stato istituito con la Decisione del Parlamento Europeo e del Consiglio n. 2481 del 14 dicembre 2022.

[6] *Ibid.* art. 4.

dell'Unione di interagire online con le amministrazioni pubbliche); il 100% dei cittadini dell'Unione potrà avere accesso al proprio fascicolo sanitario elettronico; il 100% dei cittadini dell'Unione potrà accedere a mezzi di identificazione elettronica sicura riconosciuti in tutta l'Unione, in modo da consentire loro il pieno controllo sulle transazioni con verifica dell'identità e dei dati personali condivisi.

È evidente che la logica sottostante agli indirizzi stabiliti è quella di una *governance* tecnologica non fondata esclusivamente sulla competitività economica ma anche sull'affermazione dei valori europei nella sfera digitale, tra cui il rispetto dei diritti fondamentali, la tutela della democrazia e la promozione della sostenibilità sociale. Altrettanto evidente è che questi traguardi saranno raggiungibili entro il prossimo lustro solo con l'utilizzo crescente di sistemi di IA sia come soluzione per la trasformazione digitale del settore produttivo sia come tecnologia di fruizione avanzata dei servizi pubblici da parte dei cittadini.

2. Dal quadro di riferimento europeo (ecosistema di eccellenza e principio di fiducia) all'AI Act

Lo straordinario valore strategico dell'IA per lo sviluppo dell'Unione Europea si evidenzia anche con l'adozione dell'*AI Act*[7], che costituisce il primo tentativo di disciplina organica dell'IA a livello sovranazionale e si configura come lo strumento

[7] Con il Regolamento (UE) 2024/1689 del 13 giugno 2024, ormai conosciuto come *AI Act*, il Parlamento e il Consiglio dell'Unione Europea hanno stabilito regole armonizzate sull'intelligenza artificiale.

normativo abilitante per la realizzazione del *Programma strategico per il decennio digitale 2030*[8].

Il Legislatore eurounitario ha, infatti, introdotto un approccio normativo, enunciando principi etici e imponendo obblighi stringenti per lo sviluppo e l'utilizzo "sicuro" dei sistemi di IA, con l'obiettivo, da un lato, di disporre strumenti a tutela della salute, della sicurezza e dei diritti dei cittadini e, dall'altro, di rafforzare la competitività del mercato dell'Unione.

[8] Il *Programma*, all'art. 3 stabilisce – tra le finalità generali – la promozione di «(…) *un ambiente digitale antropocentrico, basato sui diritti fondamentali, inclusivo, trasparente e aperto, in cui tecnologie e servizi digitali sicuri e interoperabili rispettino e rafforzino i principi, i diritti e i valori dell'Unione e siano accessibili a tutti, ovunque nell'Unione* (…), il rafforzamento della (…) *resilienza collettiva* (…), la garanzia della (…) *sovranità digitale dell'Unione in modo aperto, in particolare mediante infrastrutture digitali e di dati sicure e accessibili* (…), la promozione della (…) *diffusione e l'uso di capacità digitali intesi a ridurre il divario digitale* (…), lo sviluppo di (…) *un ecosistema globale e sostenibile di infrastrutture digitali interoperabili* (…), la promozione di (…) *un contesto normativo dell'Unione in materia digitale per sostenere la capacità delle imprese dell'Unione, soprattutto delle PMI* (…), la garanzia secondo cui (…) *la partecipazione online alla vita democratica sia possibile per tutti e che anche i servizi pubblici e i servizi sanitari e di assistenza siano accessibili a tutti nel quadro di un ambiente online fidato e sicuro, in particolare ai gruppi svantaggiati, comprese le persone con disabilità, e in zone rurali e remote, offrendo servizi e strumenti inclusivi, efficienti, interoperabili e personalizzati con standard elevati in materia di sicurezza e privacy* (…), la garanzia secondo cui (…)*le infrastrutture e le tecnologie digitali, comprese le relative catene di approvvigionamento, diventino più sostenibili, resilienti ed efficienti sotto il profilo energetico e delle risorse* (…), condizioni (…)*eque e non discriminatorie per gli utenti durante la trasformazione digitale in tutta l'Unione* (…), la garanzia secondo cui (…) *tutte le politiche e i programmi pertinenti ai fini del conseguimento degli obiettivi digitali* (…) *siano presi in considerazione in modo coordinato e coerente per contribuire pienamente alla transizione verde e digitale, evitando nel contempo le sovrapposizioni e riducendo al minimo gli oneri amministrativi* (…), il miglioramento della (…)*resilienza agli attacchi informatici* (…) *una maggiore consapevolezza dei rischi e una migliore conoscenza dei processi di cibersicurezza, intensificando gli sforzi delle organizzazioni pubbliche e private* (…)».

A differenza degli Stati Uniti, indirizzati a una minima regolazione delle norme sull'IA per favorire il mercato, e della Cina, interessata a una pianificazione centralizzata dei sistemi di IA, l'Unione Europea ha realizzato uno strumento regolatorio orientato a realizzare una piena sovranità digitale sulla materia, che bilancia la competitività economica con la salvaguardia della sicurezza nazionale e dei diritti fondamentali[9]. Questo approccio, affermato sin dal *Libro bianco*, ha influito sugli indirizzi normativi successivi, orientandoli alla promozione dell'utilizzo dell'IA in una prospettiva di riduzione dei rischi a questo connessi.

Dato il modello di "regolazione *ex ante*" adottato con l'*AI Act*, incentrato sulla valutazione del rischio connesso all'utilizzo dei sistemi di IA (rischio inaccettabile, alto rischio, rischio limitato e rischio minimo)[10], con le conseguenti prescrizioni e cautele a tutela dei diritti dei cittadini, è possibile cogliere una serie di connessioni tra le previsioni programmatiche e quelle regolamentari dell'*AI Act*. Infatti, sia il *Programma strategico per il decennio digitale 2030* che l'*AI Act* promuovono le competenze digitali, così come il riferimento a infrastrutture digitali sicure, alla digitalizzazione delle PMI e all'utilizzo dei sistemi di IA, che costituiscono elementi di connessione tra le norme regolamentari e gli obiettivi stabiliti dal

[9] G. FERRARI, P. DIANA, *I tentativi di regolamentazione dell'intelligenza artificiale: tra approcci differenti, contraddizioni e implicazioni sociali in Europa, Stati Uniti e Cina*, in «Sicurezza e Scienze Sociali», 2, 2025, 143-155; C. PERLINGIERI, *Intelligenza artificiale tra princìpi e regole*, in «Ai Law - International Review of Artificial Intelligence Law», 1, 2025, 40–41. Per un ulteriore approfondimento in chiave comparativistica, *cfr.* A. STIANO, *La corsa alla regolamentazione dell'intelligenza artificiale; una riflessione in chiave comparata*, in «Diritto pubblico comparato ed europeo», 1, 2025, 97–116.

[10] Per un approfondimento sul "modello di regolazione precauzionale" fornito dall'*AI Act cfr.* W. GASPARRI, F. TESI, *Intelligenza artificiale e "AI Act". Il perimetro della difficile sussunzione del soggetto nella macchina calcolante*, in «Diritto pubblico», 2, 2025, 517–524.

Programma, che l'*AI Act* contribuisce a raggiungere con la tutela dei diritti fondamentali e l'affermazione di un modello europeo di innovazione digitale, stabilendo l'equilibrio tra sovranità tecnologica, competitività economica e primato del diritto.

I riflessi di questa regolamentazione sulle attività della Pubblica Amministrazione sono assolutamente rilevanti. L'*AI Act*, nella sua valenza normativa e politica impone a tutti gli utilizzatori dei sistemi di IA un approccio basato sul rischio, che costituisce un modello più rigido rispetto ad altri, come quello limitato alla responsabilità. Il principio del "rischio", infatti, appare più idoneo ad accompagnare lo sviluppo rapidissimo dei sistemi di IA, imponendo il bilanciamento costante delle esigenze di innovazione tecnologica con il rispetto dei diritti fondamentali dei cittadini dell'Unione, rispetto a quello di mera "responsabilità", che porrebbe enfasi sulla responsabilità legale di sviluppatori e utenti. Tale impostazione, tuttavia, anche in presenza di adeguati meccanismi di controllo e di rigorosi meccanismi sanzionatori, lascerebbe agli operatori del settore margini d'azione ben più ampi, che troverebbero contenimento soltanto dopo l'accertamento di violazioni[11].

L'*AI Act*, per quanto stabilisca che l'approccio basato sul rischio costituisce la base per un insieme proporzionato ed efficace di regole vincolanti, non ha tralasciato un riferimento ai sette principi etici enunciati dal Gruppo di Esperti di Alto Livello sull'Intelligenza Artificiale (AI HLEG) con le *Linee Guida per un'IA affidabile* del 2019 e, riprendendone gli orientamenti nel "considerato" n. 27 del testo[12], ha enunciato l'altro principio fon-

[11] G. ZICCARDI, *Una lettura dell'Artificial Intelligence Act: norme, etica, adempimenti, attuazione*, in F. BASILE, M. BIASI, L. CAMALDO, G. CANESCHI, B. FRAGASSO, D. MILANI (a cura di), *Intelligenza Artificiale. Diritto, giustizia, economia ed etica*, Giappichelli, Torino, 2025, 18.

[12] «*Sebbene l'approccio basato sul rischio costituisca la base per un insieme proporzionato ed efficace di regole vincolanti, è importante ricordare gli orientamenti*

dativo della regolamentazione eurounitaria, quello antropocentrico, secondo cui «(…) *tali orientamenti contribuiscono all'elaborazione di un'IA coerente, affidabile e antropocentrica, in linea con la Carta e con i valori su cui si fonda l'Unione* (…)», che è di straordinaria pregnanza per ogni riflessione sull'utilizzo dei sistemi di IA nelle amministrazioni pubbliche.

Il principio antropocentrico, infatti, interviene per ribadire la strumentalità dell'IA rispetto ai "bisogni umani" e, soprattutto, l'imprescindibilità del controllo dei risultati algoritmici, che non potranno mai essere frutto di una delega incondizionata da parte degli esseri umani[13].

Se il principio antropocentrico non permeasse l'intera struttura regolatoria dell'*AI Act*, potrebbe concretizzarsi il rischio di un utilizzo dell'IA in ambito di lavoro pubblico, in cui si potrebbe fare affidamento sui risultati definitivi "decisi" dagli algoritmi non solo per le attività di natura vincolata ma anche per l'attività discrezionale, con la possibilità che questo tipo di attività possa compromettere l'intera architettura normativa della tutela dei diritti.

etici per un'IA affidabile del 2019 elaborati dall'AI HLEG indipendente nominato dalla Commissione. In tali orientamenti l'AI HLEG ha elaborato sette principi etici non vincolanti per l'IA che sono intesi a contribuire a garantire che l'IA sia affidabile ed eticamente valida. I sette principi comprendono: intervento e sorveglianza umani, robustezza tecnica e sicurezza, vita privata e governance dei dati, trasparenza, diversità, non discriminazione ed equità, benessere sociale e ambientale e responsabilità (…)» (G. ZICCARDI, *Una lettura dell'Artificial Intelligence Act: norme, etica, adempimenti, attuazione*, *cit.*, 21 ss.).

[13] P. FORTE, *Elementi di Intelligenza Amministrativa Artificiale (A-AI). Su taluni caratteri giuridici dell'automazione della pubblica amministrazione*, in «Diritto pubblico», 2, 2025, 415–416.

3. Il principio antropocentrico e la Legge 132 del 2025 sull'Intelligenza artificiale

Il principio antropocentrico, così come delineato nell'*AI Act*, assume il valore di vero e proprio paradigma e, come tale, è recepito dal Legislatore nazionale con l'approvazione della Legge n. 132 del 23 settembre 2025, che segna un vero e proprio momento di svolta nel panorama normativo italiano, ponendosi come quadro regolatorio di accompagnamento alla disciplina dell'*AI Act*.

L'obiettivo dichiarato delle norme stabilite dalla Legge n. 132 del 2025 è quello di porre l'autodeterminazione umana al centro di ogni fase di sviluppo e di utilizzo dei sistemi di IA, tutelando l'autonomia e il potere decisionale dell'individuo.

In questa prospettiva, l'architettura normativa proposta dalle norme nazionali gerarchizza diversamente i principi di riferimento rispetto all'*AI Act*. A differenza di quanto stabilito dal Legislatore eurounitario, che propone regole riferite al rischio connesso all'utilizzo dei sistemi di IA, quello nazionale incentra le premesse della disciplina sulla promozione dei sistemi di IA secondo i principi di un utilizzo corretto, trasparente e responsabile, radicandola saldamente in una dimensione antropocentrica, garantendo la vigilanza sui rischi economici e sociali e sull'impatto sui diritti fondamentali, secondo quanto disposto all'art. 1 del testo condiviso dalle Camere[14].

In questo senso, il comma 1 dell'articolo 1 della Legge stabilisce che l'IA deve essere impiegata per migliorare le condizioni di lavoro, tutelare l'integrità psico-fisica dei lavoratori e accrescere la qualità delle prestazioni lavorative e la produttività delle persone[15].

Le norme, quindi, stabiliscono che l'uso dell'IA in ambito la-

[14] Legge 23 settembre 2025 n. 132, art. 1, comma 1.

[15] *Ibid.*, art. 11, comma 1.

vorativo è corretto quando sono soddisfatte queste tre condizioni, a cui si aggiunge l'obbligo che i sistemi di IA utilizzati siano sicuri, affidabili e trasparenti.

Tenuto conto anche del rimando all'utilizzo dei sistemi di IA in conformità al diritto dell'Unione Europea, l'art. 11 della Legge sancisce il divieto assoluto di utilizzo quando questo sia in contrasto con la dignità umana o violi il diritto alla riservatezza. Quanto alla gestione e all'organizzazione dei rapporti di lavoro, il Legislatore italiano impone un'ulteriore assicurazione: il rispetto dei diritti inviolabili del lavoratore «(…) *senza discriminazioni in funzione del sesso, dell'età, delle origini etniche, del credo religioso, dell'orientamento sessuale, delle opinioni politiche e delle condizioni personali, sociali ed economiche, in conformità al diritto dell'Unione europea* (…)»[16].

Con riferimento alle prescrizioni normative, è interessante notare come la proposta di incrementare con i sistemi di IA la qualità delle prestazioni lavorative non possa avvenire a discapito dell'integrità psico-fisica e della dignità dei lavoratori e come tale impostazione sia la premessa delle successive prescrizioni dell'art. 14 della Legge n. 132, che si rivolge specificamente all'uso dell'IA nella Pubblica Amministrazione.

In questo caso, riferendosi alle amministrazioni pubbliche, il Legislatore stabilisce che lo scopo dell'utilizzo dei sistemi di IA è esclusivamente indirizzato all'incremento dell'efficienza amministrativa e alla riduzione dei tempi procedimentali, assicurando altresì che l'uso dell'IA è sottoposto a due criteri di garanzia: la conoscibilità del funzionamento e la tracciabilità dell'utilizzo per gli interessati.

Dopo questa dichiarazione di principio e l'enunciazione dei criteri di garanzia, il Legislatore fa proprio il principio antropocen-

[16] *Ibid.*, art. 11 comma 3.

trico recepito in ambito eurounitario e lo declina, accordandolo all'impostazione della Legge 132, secondo un'indicazione ancora più pregnante, dal punto di vista giuridico. Infatti, dopo aver riaffermato la funzione strumentale e di supporto dell'IA rispetto all'attività provvedimentale, il Legislatore riprende l'impostazione antropocentrica della disciplina, stabilendo che la «(...) *persona che resta l'unica responsabile dei provvedimenti e dei procedimenti in cui sia stata utilizzata l'intelligenza artificiale* (…)»[17].

È significativo lo spostamento terminologico operato dal Legislatore. L'articolo 14, infatti, non si rivolge al "lavoratore" ma alla "persona", intesa come soggetto titolare di diritti preesistenti a qualsiasi rapporto contrattuale di pubblico impiego, ed è la persona – non il lavoratore in quanto tale – ad essere responsabile dei procedimenti e dei provvedimenti adottati, tutelando i principi di autonomia decisionale e di responsabilità giuridica individuale oltre il perimetro dei contratti di lavoro collettivi e individuali.

In questa prospettiva di sviluppo, tuttavia, le amministrazioni pubbliche devono fare i conti con note criticità, tra le quali una grave carenza di competenze digitali del personale. Secondo uno studio del The European House Ambrosetti[18] soltanto il 5% dei circa 3 milioni e duecentomila dipendenti pubblici è in possesso di competenze adeguate a gestire l'impatto dei sistemi di IA nei processi lavorativi.

Se a questa nota e perdurante carenza, malgrado le pluriennali campagne di alfabetizzazione digitale[19], si aggiungono il rallentato

[17] *Ibid.*, art. 14, comma 2.

[18] Cfr. il documento *Le opzioni tecnologiche per la digitalizzazione avanzata della Pubblica Amministrazione*, diffuso nell'ottobre 2023 e consultato il 25 gennaio 2026 sul sito del think tank italiano The European House Ambrosetti Group, organizzatore del "Forum di Cernobbio".

[19] Nelle politiche di sviluppo delle amministrazioni pubbliche, la formazione del personale è sempre più centrale; si pensi agli obblighi formativi introdotti dal-

turnover del personale, che risulta sempre più anziano, e la riduzione della spesa in ambito ICT, le aspettative di un generale miglioramento dell'efficienza della Pubblica Amministrazione anche con l'utilizzo dei sistemi di IA potrebbe risultare piuttosto ridimensionata.

4. Le Linee Guida per l'adozione di IA nella pubblica amministrazione dell'Agenzia per l'Italia Digitale (AGID)

A intervenire in materia di competenze da sviluppare in ambito di amministrazioni pubbliche è l'Agenzia per l'Italia Digitale (AGID)[20], che, con il documento *Strategia italiana per l'Intelligenza Artificiale 2024-2026* (ai sensi del *Piano triennale per l'informatica nella pubblica amministrazione 2024-2026*)[21], ha individuato, tra le aree d'intervento strategico, anche la Pubblica Amministrazione. Gli obiettivi stabiliti dall'AGID consistono nel supportare i processi e favorire la fruizione dei servizi attraverso le tecnologie di IA, con la promozione di sei specifiche azioni strategiche, che si realizzano con *Linee guida per l'adozione di IA*, *Linee guida per il*

la cosiddetta Direttiva del Ministro della Pubblica Amministrazione del 14 gennaio 2025 *Valorizzazione delle persone e produzione di valore pubblico attraverso la formazione. Principi, obiettivi e strumenti*, con cui sono stati ampliati gli obiettivi per la formazione e lo sviluppo delle competenze dei dipendenti pubblici nelle aree della transizione digitale, ecologica e amministrativa.

[20] Ai sensi dell'art. 20, comma 1 della Legge n. 132 del 2025, l'Agenzia per l'Italia Digitale è una delle due Autorità nazionali ad applicare e ad attuare le norme nazionali ed eurounitarie in ambito di IA; la seconda Autorità nazionale è l'Agenzia per la Cybersicurezza nazionale, *cfr.* Legge n. 132 del 2025, art. 20, comma 1, lettera b).

[21] I documenti *Strategia italiana per l'Intelligenza Artificiale 2024-2026* e *Piano triennale per l'informatica nella pubblica amministrazione 2024-2026*, alla data del 20 settembre 2026, sono consultabili sul sito dell'AGID.

procurement, *Linee guida per la realizzazione di applicazioni di IA*, *Semplificazione per cittadini e imprese*, l'*Efficientamento* e l'*Introduzione dell'IA nelle scuole per la Pubblica Amministrazione*.

Delle sei azioni strategiche, l'AGID ha avviato solo la prima, il 18 febbraio 2025, con la diffusione per la consultazione pubblica delle bozze delle *Linee Guida per l'adozione di IA nella pubblica amministrazione*[22], con cui l'Agenzia, nel definire gli indirizzi per l'adozione, l'acquisto e lo sviluppo di sistemi di IA nella Pubblica Amministrazione, ha inteso raccogliere dagli *stakeholder* commenti, suggerimenti e proposte di modifica per la successiva adozione, che è attesa per il 2026.

Il documento, nell'analizzare il contesto e le caratteristiche, evidenzia – tra le altre – la necessità di rafforzare le competenze dei dipendenti[23], che secondo lo studio costituiscono uno dei principali fattori abilitanti per cogliere in "modo consapevole" i vantaggi offerti dall'utilizzo dei sistemi di Intelligenza artificiale negli ambiti del lavoro pubblico.

Il rafforzamento delle competenze dei dipendenti emerge, in definitiva, come il fattore abilitante primario per garantire che il passaggio all'IA avvenga in modo "consapevole" e conforme ai valori di dignità e trasparenza promossi dalla Legge n. 132 del 2025, nel quadro proposto dall'*AI Act.*

In questo senso le Linee Guida promuovono lo sviluppo di nuove competenze digitali, affiancando all'alfabetizzazione di base per tutti i dipendenti anche la creazione di nuovi profili specialistici, con particolare attenzione alla formazione continua e alla colla-

[22] Il documento *Linee Guida per l'adozione di IA nella pubblica amministrazione*, ancora non definitivo alla data del 20 settembre 2026, è consultabile sul sito dell'AGID.

[23] *Ibid.*, 43-50.

borazione con le università e i centri di ricerca per attrarre giovani talenti.

La Pubblica Amministrazione, quindi, deve agire sui fattori abilitanti per trasformare il potenziale algoritmico in efficienza operativa, garantendo che il capitale umano sia il motore e non il freno della transizione digitale. La formazione, quindi, non può e non deve rappresentare un corollario di adempimenti ministeriali ma lo strumento per comprendere le potenzialità dei sistemi di IA non solo nelle attività più ripetitive ma anche in quelle in cui possono esprimersi i risultati dell'analisi predittiva a supporto dei processi decisionali più complessi.

In questa prospettiva, la Pubblica Amministrazione deve garantire la promozione di una vera e propria *AI Awareness* (intesa come comprensione del funzionamento dell'IA, conoscenza delle norme in materia e consapevolezza degli impatti dell'IA sui diritti fondamentali, delle potenziali discriminazioni e delle implicazioni socio-economiche del suo utilizzo) non solo presso i dipendenti pubblici ma anche i fruitori dei servizi erogati dalla Pubblica Amministrazione, in modo da favorire la costruzione di un rinnovato rapporto di fiducia tra gli Enti e i cittadini.

La comprensione di quanto le competenze costituiscano una leva straordinaria per generare valore pubblico consente di agire su più livelli e di abilitare le amministrazioni pubbliche a strutturare la formazione dei dipendenti non solo in termini di *AI Literacy*[24] ma anche di sviluppo di competenze specialistiche sino ad oggi

[24] Il punto 56 dell'art 3 dell'*AI Act*, *cit.,* esplicita il concetto di *AI Literacy* come «(…) *alfabetizzazione in materia di IA: le competenze, le conoscenze e la comprensione che consentono ai fornitori, ai deployer e alle persone interessate, tenendo conto dei loro rispettivi diritti e obblighi nel contesto del presente regolamento, di procedere a una diffusione informata dei sistemi di IA, nonché di acquisire consapevolezza in merito alle opportunità e ai rischi dell'IA e ai possibili danni che essa può causare* (…)».

sconosciute. Gli ambiti di interesse sono la qualificazione di esperti in progettazione, di specialisti di sviluppo e gestione dei dati (*data engineer*), la formazione di esperti di algoritmi di ML (*machine learning engineer*), di specialisti per l'istruzione dei sistemi di IA in termini di efficacia dei risultati (*prompt engineer*), di specialisti degli impatti sociali dei sistemi di IA (*AI ethicist*)[25].

Queste figure specialistiche non esauriscono le competenze necessarie all'integrazione dei sistemi di IA nei processi delle amministrazioni pubbliche. Infatti, le Linee Guida inseriscono tra le figure professionali critiche non solo giuristi informatici, esperti di protezione dei dati personali, *change manager* (esperti che valutano gli impatti dell'IA ed elaborano strategie di gestione) ma anche esperti umanisti, cioè specialisti in grado di supportare tutte le altre figure su temi e questioni etiche, sociali, culturali e filosofiche.

La trasversalità delle competenze necessarie all'utilizzo consapevole dei sistemi di IA appare di particolare pregnanza per quanto riguarda le funzioni affidate ai dirigenti pubblici, che devono affiancare alle competenze di base una visione strategica. Ai dirigenti, infatti, è richiesto di valutare l'impatto, la sostenibilità etica e la conformità legale dei sistemi di IA e, per questo, devono sviluppare sia *soft skills* che capacità di *leadership*, entrambe necessarie a guidare i processi di transizione e a gestire eventuali crisi o malfunzionamenti dei sistemi.

Affinché questi obiettivi formativi siano raggiunti la Pubblica Amministrazione deve adottare una *roadmap* operativa rigorosa, in grado di includere strumenti di pianificazione come il Piano Integrato di Attività e Organizzazione (PIAO), la fruizione di corsi tematici attraverso la piattaforma Syllabus, o il Formez, e il ricorso a percorsi di formazione specialistica e d'eccellenza, offerti dalla

[25] Cfr. *Linee Guida per l'adozione di IA nella pubblica amministrazione, cit.*, 45-46.

Scuola Nazionale dell'Amministrazione (SNA) o da programmi universitari come *PA 110 e Lode*.

La valorizzazione e la formazione delle competenze dei pubblici dipendenti e, quindi, la comprensione di come e perché i sistemi di IA esplichino una tale influenza sui processi delle amministrazioni pubbliche consentono di acquisire padronanza degli aspetti normativi, operativi e di dominio associati ai sistemi di IA, a cui si aggiunge la comprensione dei principi etici legati al loro uso, incluso quello fondamentale sulla trasparenza nell'ambito di decisioni prese in maniera automatizzata.

In conclusione, la trasformazione indotta dall'utilizzo dei sistemi di Intelligenza artificiale richiede alla Pubblica Amministrazione l'adozione di una cultura dell'apprendimento continuo come pilastro fondamentale del lavoro pubblico. Solo attraverso una formazione costante, strutturata e normativamente allineata, la Pubblica Amministrazione potrà configurarsi come un ecosistema resiliente, capace di generare valore pubblico e di guidare con responsabilità la società nel futuro digitale.

EPR tessile e IA: profili di raccordo alla luce dell'AI Act

di *Carmen Parente*[*]

SOMMARIO: 1. Introduzione. – 2. Algoritmo e procedimento amministrativo: trasformazioni e criticità sistemiche. – 3. Trasparenza e motivazione nell'amministrazione algoritmica. – 4. Imputazione della decisione e responsabilità nell'amministrazione algoritmica. – 5. L'EPR tessile come laboratorio di amministrazione algoritmica. – 6. Conclusioni.

1. Introduzione

L'adozione del Regolamento UE 2024/1689 sull'Intelligenza Artificiale (AI Act)[1] segna l'istituzionalizzazione dei sistemi algoritmici nell'ordinamento europeo ponendosi come fattore strutturale capace di trasformare profondamente l'esercizio delle funzioni pubbliche e l'assetto delle garanzie[2]. In questa prospettiva, l'analisi giuridica deve verificare la resilienza dei paradigmi classici di legalità, responsabilità e tutela[3], poiché l'uso dell'Intelligenza Artificiale (IA) a supporto o in sostituzione delle valutazioni umane incide direttamente sui principi fondanti dell'azione amministrativa.

[*] Dottoranda in Società in mutamento: politiche, diritti e sicurezza, Università degli Studi della Tuscia.

[1] Regolamento UE n.1689 del 13/06/2024, Parlamento europeo e del Consiglio recante norme armonizzate sull'intelligenza artificiale (Artificial Intelligence Act), GUUE L 2024/1689, 12 luglio 2024.

[2] G. NAPOLITANO, *Amministrazione e algoritmi*, in «Giornale di diritto amministrativo», 7, 2020, 801–809.

[3] S. CASSESE, *Lo spazio giuridico globale*, Laterza, Roma-Bari, 2003, 171 ss.

La transizione verso modelli decisionali mediati dalla tecnologia fa emergere criticità sistemiche in ordine alla trasparenza e alla controllabilità del potere pubblico.

In particolare, la mediazione algoritmica tende a travalicare la dimensione meramente strumentale per incidere sulla struttura stessa del processo decisionale; in questa prospettiva, l'algoritmo non rappresenta un mero ausilio materiale, ma assume la veste giuridica di "atto a elaborazione elettronica", al quale sono affidati l'interrelazione tra norme e dati e lo svolgimento dell'iter logico-motivazionale. Tale mutamento impone di ripensare in chiave sistematica le categorie tradizionali del diritto amministrativo. In questo senso, le politiche ambientali e, in particolare, i meccanismi di Responsabilità Estesa del Produttore (Extended Producer Responsibility, EPR) offrono un osservatorio privilegiato, configurandosi come una vera e propria infrastruttura di governance multilivello[4] all'interno della quale l'Intelligenza Artificiale assume un ruolo centrale nei processi di *policy implementation*. Il settore dell'industria tessile rappresenta un ambito particolarmente fertile per l'applicazione di tali logiche, trattandosi di una filiera dal doppio volto: caratterizzata da un elevato impatto ambientale ma con ampi margini di circolarità ancora non valorizzati.

Uno degli strumenti chiave di questa transizione è il Passaporto Digitale dei Prodotti (Digital Product Passport, DPP), istituito dal Regolamento (UE) 2024/1781[5], attraverso il quale l'IA permette di gestire obblighi informativi vincolanti su composizione del materiale, tracciabilità e riparabilità, consentendo di presidiare

[4] E. CHITI, *Economia circolare e diritto pubblico*, in «Rivista trimestrale di diritto pubblico», 2021, 789 ss.

[5] Regolamento UE n. 1781 del 13/06/2024, Parlamento europeo e del Consiglio che istituisce un quadro per la definizione dei requisiti di progettazione ecocompatibile (Ecodesign for Sustainable Products Regulation) per prodotti sostenibili e che abroga la direttiva 2009/125/CE, GUUE L 2024/1781.

l'intero ciclo di vita del prodotto, dalla progettazione iniziale fino all'internalizzazione dei costi di fine vita.

La progressiva introduzione di strumenti quali il *Digital Product Passport*, i sistemi di tracciabilità lungo la catena del valore e i meccanismi di condivisione dei dati previsti dalla disciplina europea configura il settore tessile come uno dei contesti nei quali si manifestano in modo particolarmente evidente le tendenze verso modelli di amministrazione *data-driven*.

Muovendo da tali premesse, il presente contributo intende analizzare il rapporto tra algoritmo e funzione pubblica, verificando la tenuta dei principi di legalità e trasparenza, affinché la trasformazione digitale dell'ecosistema tessile resti ancorata alle garanzie dell'ordinamento europeo e contribuisca al raggiungimento dei traguardi di circolarità fissati per il 2030 adottata dalla Commissione europea nel 2022.

2. Algoritmo e procedimento amministrativo: trasformazioni e criticità sistemiche

Nel modello tradizionale del diritto amministrativo, il procedimento assolve a una funzione di razionalizzazione e legittimazione della decisione, consentendo la riconducibilità dell'esercizio del potere a un soggetto pubblico responsabile e rendendo possibile il controllo, in sede amministrativa e giurisdizionale, sulla conformità dell'azione ai principi costituzionali, sanciti dall'art. 97 Cost., di legalità, imparzialità e buon andamento[6]. In questa prospettiva classica, il procedimento rappresenta lo strumento attraverso cui il

[6] M. CLARICH, *Manuale di diritto amministrativo*, il Mulino, Bologna, 2023, spec. cap. I.

potere pubblico si rende giuridicamente controllabile, trasformando l'autorità in attività giuridicamente organizzata[7].

L'integrazione di sistemi algoritmici nell'azione amministrativa impone una riconsiderazione del rapporto tra tecnologia e procedimento, inteso non solo come sequenza di atti funzionali all'adozione della decisione, ma come struttura portante di garanzia dell'esercizio del potere pubblico[8].

Le prime riflessioni sull'impatto delle tecnologie informatiche sull'azione amministrativa si sono sviluppate in una prospettiva prevalentemente strumentale, volta a verificare la compatibilità dell'utilizzo di strumenti elettronici con le forme tradizionali del procedimento. In tale fase, l'innovazione tecnologica veniva concepita come fattore idoneo a incidere sull'efficienza organizzativa e sull'automazione di attività materiali o istruttorie, senza mettere in discussione la struttura sostanziale del processo decisionale amministrativo[9]. La tecnologia si collocava, dunque, in una posizione esterna rispetto al nucleo valutativo e decisorio dell'azione amministrativa.

Il progressivo ricorso a sistemi algoritmici, e in particolare a modelli di Intelligenza Artificiale, segna un mutamento qualitativo: l'algoritmo non si limita più a eseguire istruzioni predeterminate, ma tende a orientare la ponderazione degli interessi e l'esito del procedimento, incidendo direttamente sulla formazione della volontà amministrativa. In questo passaggio, il contributo tecnologico assume un carattere determinante o preponderante, condizio-

[7] M.S. GIANNINI, *Istituzioni di diritto amministrativo*, 2ª ed., Giuffrè, Milano, 2000, cap. I.

[8] A. MASUCCI, *Procedimento amministrativo e nuove tecnologie*, Giappichelli, Torino, 2011, 89 ss.

[9] G. DUNI, *L'utilizzabilità delle tecniche elettroniche nell'emanazione degli atti e nei procedimenti amministrativi*, in «Rivista amministrativa della Repubblica italiana», 1978, 407 ss.

nando la decisione finale nella sua dimensione sostanziale e rischiando di alterare il tradizionale nesso tra discrezionalità e responsabilità, introducendo una distanza crescente tra la formalità del procedimento e la sostanza del processo decisionale.

Questa dinamica pone interrogativi di natura sistemica sulla tenuta del modello procedimentale. Il procedimento amministrativo presuppone, infatti, la possibilità di ricostruire il percorso logico-giuridico che conduce alla decisione attraverso fasi formalizzate, obblighi di partecipazione e oneri motivazionali. I sistemi algoritmici, soprattutto quando fondati su modelli complessi e adattivi, operano invece secondo logiche che possono rendere opaca la correlazione tra dati di input, criteri di elaborazione e risultato finale. Il rischio che ne deriva è quello di una progressiva scissione tra procedimento formale e decisione sostanziale, con una traslazione del momento decisorio verso ambiti tecnici sottratti alle ordinarie garanzie del diritto amministrativo[10].

Le criticità evidenziate risultano particolarmente accentuate nei settori caratterizzati da elevata complessità tecnica e da decisioni seriali fondate sull'elaborazione di grandi quantità di dati, come quello ambientale e, in particolare, i sistemi di Responsabilità Estesa del Produttore (EPR). In tali ambiti, l'algoritmo non si limita a supportare il procedimento, ma ne struttura l'andamento complessivo, incidendo sulla determinazione degli obblighi giuridici dei destinatari e sulla ripartizione dei costi[11]. Tale dinamica emerge con particolare chiarezza nei regimi EPR applicati al settore tessile, nei quali la quantificazione dei contributi a carico dei produttori risulta strettamente connessa alla raccolta e all'elaborazione di dati relativi alle quantità immesse sul mercato, alla com-

[10] Consiglio di Stato, sez. VI, 8 aprile 2019, n. 2270.

[11] F. MUNARI, *Responsabilità estesa del produttore e governance ambientale*, in «Ambiente e diritto», 2020.

posizione dei prodotti e alle performance ambientali, secondo un modello fortemente *data-driven* delineato dalla normativa europea e dal processo di recepimento nazionale.

3. Trasparenza e motivazione nell'amministrazione algoritmica

L'introduzione di sistemi algoritmici nei procedimenti amministrativi incide in modo diretto su due principi cardine dell'azione amministrativa: la trasparenza e la motivazione. Entrambi costituiscono presupposti essenziali della legittimità dell'azione pubblica, in quanto funzionali non solo alla tutela delle posizioni giuridiche dei destinatari, ma anche all'effettività del controllo sull'esercizio del potere amministrativo. Nel modello tradizionale, trasparenza e motivazione consentono di ricostruire il percorso logico-giuridico seguito dall'amministrazione, rendendo intelligibili le ragioni della decisione e assicurandone la sindacabilità in sede giurisdizionale[12].

L'impiego di algoritmi decisionali introduce, tuttavia, elementi di significativa discontinuità rispetto a tale assetto. I sistemi di Intelligenza Artificiale, soprattutto quando basati su tecniche di apprendimento automatico (*machine learning*), incluse le applicazioni fondate su modelli di tipo profondo (*deep learning*), presentano profili di particolare criticità sotto il profilo della trasparenza e della motivazione. Essi operano attraverso modelli decisionali caratterizzati da un elevato grado di complessità tecnica e, in molti casi, da una strutturale opacità, che rende difficile la ricostruzione del nesso tra dati di input, criteri di elaborazione e risultato finale[13]. In tali ipotesi, la concezione tradizionale della motivazione come

[12] Consiglio di Stato, sez. IV, 3 giugno 2019, n. 3692.

[13] G. NAPOLITANO, cit.

esposizione delle ragioni giuridiche e fattuali della decisione risulta messa in crisi.

La giurisprudenza amministrativa ha progressivamente chiarito che il ricorso a strumenti algoritmici non può tradursi in una compressione delle garanzie procedimentali, ribadendo che l'amministrazione resta responsabile della decisione adottata e tenuta a renderne intelligibili le modalità di formazione[14]. In questa prospettiva, l'algoritmo non può essere considerato una "scatola nera" sottratta al controllo giuridico, ma deve essere assoggettato a forme di trasparenza idonee a consentire la verifica della correttezza, coerenza e non arbitrarietà del procedimento decisionale[15].

La trasparenza dell'azione amministrativa algoritmica non si esaurisce, tuttavia, nella mera accessibilità del codice sorgente o nella descrizione tecnica del funzionamento del sistema. Essa richiede piuttosto la possibilità di comprendere il ruolo effettivamente svolto dall'algoritmo nel procedimento e il grado di incidenza esercitato sulla decisione finale. Come evidenziato in dottrina, la trasparenza deve essere declinata in senso funzionale, quale conoscibilità delle regole decisionali incorporate nel sistema e delle scelte discrezionali operate dall'amministrazione nelle fasi di progettazione, addestramento e implementazione dell'algoritmo[16].

Analoghe considerazioni valgono con riferimento all'obbligo di motivazione.

Quando la decisione amministrativa è il risultato di un processo decisionale mediato dall'algoritmo, la motivazione non può limitarsi a richiamare l'esito dell'elaborazione automatizzata, ma deve

[14] Consiglio di Stato, sez. VI, 13 dicembre 2019, n. 8472.

[15] TAR Lazio, sez. III-bis, 22 marzo 2019, n. 3769.

[16] D. PALAZZO, *Intelligenza artificiale e scelte amministrative. La discrezionalità alla prova degli algoritmi*, in *Dialoghi di diritto amministrativo. Lavori del laboratorio di diritto amministrativo* 2022-2023, a cura di F. Aperio Bella, A. Carbone, E. Zampetti, Roma TRE Press, Roma, 2024, 91–146.

dar conto delle ragioni per cui l'amministrazione ha fatto ricorso a tale strumento e delle modalità con cui esso è stato utilizzato[17]. Ciò implica che l'obbligo motivazionale si estenda anche alle scelte a monte del procedimento, relative alla selezione dei dati, alla definizione dei parametri decisionali e alla configurazione del sistema algoritmico.

Il problema si pone in termini particolarmente delicati nei casi di decisioni seriali e standardizzate, nei quali il rischio di una riduzione della motivazione a formule stereotipate risulta accentuato. In tali ipotesi, l'automazione può tradursi in una motivazione meramente apparente, incapace di assolvere alla funzione di garanzia che l'ordinamento le attribuisce[18]. Ne deriva l'esigenza di ripensare il contenuto e le modalità della motivazione amministrativa alla luce delle nuove forme di decisione algoritmica, evitando che le esigenze di efficienza e rapidità si traducano in una compressione delle garanzie sostanziali dei destinatari[19].

Nel contesto delle politiche ambientali e, in particolare, dei sistemi di Responsabilità Estesa del Produttore, le criticità legate all'amministrazione algoritmica assumono un rilievo specifico. L'utilizzo di sistemi basati su algoritmi per la tracciabilità dei prodotti, il monitoraggio dei flussi di rifiuti e la determinazione degli oneri economici comporta l'adozione di decisioni amministrative fondate su elaborazioni complesse e valutazioni tecniche. In tale contesto, il Passaporto Digitale dei Prodotti (DPP) garantisce la disponibilità di dati strutturati – quali la composizione fibrosa e la riparabilità – che l'algoritmo deve elaborare in modo trasparente per definire la responsabilità finanziaria del produttore.

Trasparenza e motivazione costituiscono strumenti essenziali

[17] M. CLARICH, cit.

[18] Consiglio di Stato, n. 2770, cit.

[19] G. NAPOLITANO, cit.

per garantire la legittimazione dell'azione amministrativa e per assicurare un controllo effettivo sulla corretta applicazione dei regimi EPR[20].

Il Regolamento (UE) 2024/1689 (AI Act) si colloca esattamente in questa prospettiva, pur non imponendo un obbligo generalizzato di spiegabilità, e introduce requisiti di trasparenza, documentazione e tracciabilità particolarmente rilevanti per le applicazioni idonee a incidere su decisioni amministrative con effetti giuridici diretti. Tali previsioni rafforzano l'esigenza di un coordinamento tra la disciplina dell'Intelligenza Artificiale e i principi del diritto amministrativo, imponendo all'amministrazione un uso consapevole e responsabile degli strumenti algoritmici. Il Regolamento rappresenta un punto di svolta tanto per l'approccio europeo alla regolazione dell'IA quanto per la definizione di standard innovativi in materia di gestione del rischio e responsabilità, offrendo spunti rilevanti per le traiettorie evolutive della disciplina e per le sfide applicative che attendono operatori e *policy maker*[21].

Alla luce di tali considerazioni, la sfida posta dall'amministrazione algoritmica non consiste tanto nell'adattare formalmente gli istituti tradizionali alle nuove tecnologie, quanto nel preservarne la funzione sostanziale. Trasparenza e motivazione devono continuare ad operare come presidi della legalità e della responsabilità dell'azione amministrativa, anche – e soprattutto – in contesti decisionali sempre più mediati dalla tecnologia[22].

[20] F. MUNARI, cit.

[21] A. SAURWEIN, *The EU Artificial Intelligence Act – A Landmark Proposal*, in «Oxford Business Law Blog», 2024.

[22] J. KROLL *et al.*, *Accountable Algorithms*, in «University of Pennsylvania Law Review», 165, 2016, 695 ss.

4. Imputazione della decisione e responsabilità nell'amministrazione algoritmica

L'automazione decisionale introduce una tensione profonda tra le logiche tradizionali di imputazione del potere amministrativo e le nuove modalità di produzione della decisione, ponendo in modo particolarmente problematico la questione della responsabilità per gli effetti giuridici che ne derivano. Nel modello tradizionale del diritto amministrativo, la decisione è imputata all'amministrazione quale titolare del potere, attraverso l'azione di organi e funzionari individuabili, cui è possibile ricondurre tanto la scelta quanto le conseguenze giuridiche dell'esercizio del potere pubblico.

L'introduzione dell'algoritmo come elemento strutturante del processo decisionale tende, tuttavia, a mettere in tensione tale schema. Nei regimi di Responsabilità Estesa del Produttore applicati al settore tessile, la determinazione degli obblighi economici dei produttori risulta sempre più affidata a processi di elaborazione algoritmica fondati su dati gestiti da una pluralità di soggetti, pubblici e privati, con conseguente frammentazione della catena decisionale e attenuazione del tradizionale nesso tra potere, decisione e responsabilità amministrativa. Quando l'esito del procedimento è il risultato di un'elaborazione automatizzata o semiautomatizzata, il contributo umano rischia di ridursi a una funzione meramente formale di validazione dell'esito prodotto dal sistema.

Come chiarito in modo concorde da dottrina e giurisprudenza, il ricorso a strumenti algoritmici non comporta alcuna deresponsabilizzazione dell'autorità procedente, la quale resta l'unica titolare del potere decisionale e responsabile degli effetti prodotti. In questa prospettiva, la natura tecnica o automatizzata del procedi-

mento non esonera l'amministrazione dall'obbligo di verificare la correttezza dell'iter seguito e di assumersi la paternità della decisione finale; tuttavia, tale principio, pur saldo sul piano formale, rischia di essere messo in crisi quando il contributo algoritmico assume carattere determinante. In tali ipotesi, l'imputazione giuridica della scelta può occultare una delega di fatto a modelli tecnici progettati secondo logiche non immediatamente riconducibili alla razionalità giuridica, rendendo complessa la reale sindacabilità delle scelte incorporate nell'algoritmo.

Il problema si accentua ulteriormente quando l'amministrazione si avvale di sistemi algoritmici sviluppati e gestiti da soggetti privati, nell'ambito di rapporti contrattuali o di partenariato pubblico-privato (PPP). In questi casi, la catena decisionale si articola su più livelli, coinvolgendo non solo l'autorità pubblica, ma anche i fornitori tecnologici responsabili della progettazione, dell'addestramento e dell'aggiornamento degli algoritmi[23]. Ne deriva un'esigenza di chiarificazione dei criteri di imputazione e di riparto delle responsabilità, nonché del perimetro del controllo pubblico sulle scelte tecniche incorporate nei sistemi utilizzati. Resta tuttavia aperto il problema della concreta sindacabilità delle scelte algoritmiche e della possibilità per il giudice di esercitare un controllo effettivo su decisioni fondate su modelli complessi e opachi.

In questo contesto, assume rilievo centrale il tema del controllo giurisdizionale sull'algoritmo. Il sindacato del giudice amministrativo, tradizionalmente orientato alla verifica della legittimità dell'atto e della correttezza del procedimento, è chiamato a confrontarsi con decisioni la cui razionalità interna non è immediatamente accessibile. Come evidenziato in dottrina, ciò impone un

[23] F. FAINI, *Il diritto nella tecnica: tecnologie emergenti e nuove forme di regolazione*, in «Federalismi.it», 2020.

adattamento delle tecniche di controllo, volto a evitare che l'automazione decisionale si traduca in una zona franca sottratta alla giurisdizione[24]. Il controllo non può pertanto arrestarsi all'atto finale, ma deve estendersi, nei limiti del sindacato di legittimità, alla verificabilità delle regole decisionali incorporate nel sistema e delle scelte effettuate dall'amministrazione nella sua adozione e nel suo impiego. In tale prospettiva, l'algoritmo non può essere considerato un mero strumento tecnico neutrale, ma deve essere ricondotto a pieno titolo nell'ordine giuridico, quale elemento idoneo a incidere sulle modalità di esercizio del potere pubblico e sui criteri di imputazione della decisione[25]. L'amministrazione è pertanto chiamata non solo ad assumere la decisione finale, ma a governare consapevolmente l'intero processo decisionale tecnologicamente mediato.

L'AI Act si inserisce in questo quadro rafforzando la dimensione dell'accountability nell'uso dei sistemi algoritmici. I requisiti di governance, documentazione e tracciabilità previsti dal Regolamento contribuiscono a rendere più ricostruibile la catena decisionale e, conseguentemente, a rendere effettiva l'attribuzione delle responsabilità anche nei procedimenti amministrativi fondati sull'Intelligenza Artificiale.

Nel settore delle politiche ambientali e, in particolare, nei sistemi di EPR, tali problematiche emergono con particolare evidenza. L'utilizzo di algoritmi per la determinazione degli obblighi economici dei produttori, per la gestione dei flussi informativi e per il monitoraggio delle performance ambientali[26] comporta l'adozione di decisioni seriali e standardizzate, spesso fondate su

[24] J. KROLL *et al.*, cit.

[25] A. STERPA, *L'ordine giuridico dell'algoritmo*, Editoriale Scientifica, Napoli, 2024.

[26] F. MUNARI, cit.

modelli tecnici complessi. In questi contesti, il rischio è quello di una responsabilità "diluita", nella quale la decisione risulta formalmente imputata all'amministrazione ma sostanzialmente determinata da logiche algoritmiche difficilmente contestabili. Ciò impone un rafforzamento dei presidi di governo pubblico dell'algoritmo, quale condizione per preservare il nesso tra potere, responsabilità e controllo.

Nel contesto dell'EPR tessile, l'affidamento crescente a sistemi algoritmici pone sfide di particolare rilievo in materia di trasparenza e accountability, soprattutto quando le decisioni amministrative sono il risultato di processi automatizzati difficilmente verificabili dall'esterno. È stato evidenziato come il ricorso all'Intelligenza Artificiale richieda non soltanto il rispetto di nuovi standard tecnici imposti dalla normativa europea, ma anche una ridefinizione della governance pubblica, affinché l'amministrazione non si limiti a una validazione formale delle decisioni automatizzate, ma mantenga il controllo e la responsabilità sostanziali sull'intero procedimento. L'adozione di meccanismi di audit algoritmico, la documentazione accurata e la partecipazione degli stakeholder risultano dunque essenziali per garantire una reale accountability e per prevenire derive opache o deresponsabilizzanti nelle scelte di interesse collettivo[27].

5. L'EPR tessile come laboratorio di amministrazione algoritmica

L'efficacia dei regimi di EPR nel settore tessile dipende in modo decisivo dall'integrazione di fattori tecnici e informativi. Come evidenziato dalla letteratura interdisciplinare, la tracciabilità dei

[27] M. VEALE, *Algorithms and Accountability in Government*, in «Regulation & Governance», 2021.

materiali, la misurazione delle performance ambientali e la gestione dei dati lungo l'intero ciclo di vita del prodotto tessile costituiscono elementi decisivi per l'effettività delle politiche di sostenibilità.

In questa prospettiva, la transizione verso modelli circolari – come il modello di *Eco-Circular Denim*[28], inteso come sistema integrato di gestione del ciclo di vita del prodotto fondato su tracciabilità dei materiali, monitoraggio ambientale e responsabilità estesa del produttore – richiede un'infrastruttura capace di connettere innovazione tecnologica e meccanismi di governance. Tale esigenza rende centrale il ricorso a strumenti digitali e sistemi algoritmici per il monitoraggio dei rifiuti e la determinazione degli obblighi economici dei produttori: in questo assetto, l'amministrazione non svolge un mero ruolo istruttorio, ma governa processi decisionali.

I sistemi EPR richiedono la raccolta e l'elaborazione di informazioni eterogenee relative alle quantità immesse sul mercato, alla composizione dei prodotti tessili, ai flussi di raccolta e trattamento dei rifiuti, nonché ai costi connessi alla gestione della fase post-consumo. L'utilizzo di strumenti algoritmici consente di gestire tale complessità, ma incide direttamente sulle modalità di esercizio della funzione amministrativa, orientando la determinazione degli obblighi finanziari e organizzativi gravanti sui produttori[29]. Ne deriva una forma di amministrazione *data-driven*, nella quale la decisione pubblica è il risultato di elaborazioni tecniche che condizionano in modo significativo l'esito del procedimento. Tale dimensione empirica rafforza la qualificazione dell'EPR tessile come am-

[28] C C. PARENTE, M. TARANTINO, A. STERPA, E. M. MOSCONI, *Eco-Circular Denim: A Systematic Literature Review on the Eco-Sustainable and Circular Urban Model*, in *Environmental Engineering and Management Journal*, vol. 24, n. 10, 2025, pp. 2195-2207.

[29] F. MUNARI, cit.

bito nel quale l'uso dell'Intelligenza Artificiale non rappresenta un'opzione eventuale, ma una componente strutturale del modello di governance.

In questo quadro, il ricorso a sistemi algoritmici non si limita a una funzione di supporto operativo, ma incide sulla struttura stessa dei rapporti tra amministrazione e operatori economici. Questo rende particolarmente delicato il tema della trasparenza delle regole decisionali incorporate nei sistemi e della possibilità, per i destinatari, di comprendere e contestare le scelte amministrative che ne derivano. L'adozione di tecnologie innovative e di best practice costituisce un fattore chiave per promuovere la sostenibilità e facilitare la transizione verso un'economia circolare nel settore tessile[30].

Tali modelli non risultano tuttavia neutrali dal punto di vista economico. La configurazione degli algoritmi utilizzati per la determinazione degli oneri EPR produce effetti redistributivi rilevanti, incidendo sulla struttura dei costi delle imprese, sulla competitività dei diversi operatori e, in ultima analisi, sull'assetto dei mercati[31]. In questo senso, l'algoritmo assume una funzione assimilabile a quella di uno strumento di regolazione economica implicita, capace di orientare comportamenti industriali e scelte di investimento, pur in assenza di decisioni formalmente discrezionali[32].

La Responsabilità estesa del Produttore nel settore tessile pone inoltre problemi specifici di coordinamento tra soggetti pubblici e privati. I sistemi collettivi di gestione, frequentemente incaricati di

[30] E. GHISELLINI *et al.*, *Circular Economy in the Textile Sector: A review of technologies, best practices and sustainability frameworks*, in «Sustainability», 2022.

[31] F. PASQUALE, *The Black Box Society: The Secret Algorithms That Control Money and Information*, Cambridge (MA), Harvard University Press, Cambridge, 2015.

[32] K. YEUNG, *Algorithmic Regulation: A critical interrogation*, in «Regulation & Governance», 12, 2018, 505–523.

operare in nome e per conto dei produttori, svolgono un ruolo centrale nella raccolta dei dati e nell'elaborazione delle informazioni necessarie all'attuazione del regime EPR. In tali assetti, l'amministrazione esercita prevalentemente funzioni di regolazione e controllo, affidandosi a infrastrutture informative e modelli tecnici sviluppati e gestiti da soggetti terzi[33], con conseguente accentuazione del rischio di una dipendenza tecnologica dell'azione amministrativa e della necessità di predisporre presidi giuridici idonei a garantire un controllo pubblico effettivo sui processi decisionali.

Tale esigenza si collega strettamente agli approfondimenti sull'evoluzione digitale del procedimento amministrativo, che evidenziano come la digitalizzazione rappresenti una leva fondamentale per assicurare trasparenza, efficienza e tracciabilità nei rapporti tra amministrazione e soggetti privati[34].

In questa prospettiva, le trasformazioni organizzative della pubblica amministrazione e l'innovazione tecnologica impongono una riflessione sulla ridefinizione dei processi e dei ruoli giuridici, soprattutto nei settori caratterizzati da elevata complessità tecnica, come quello tessile[35].

L'intreccio tra innovazione digitale e *policymaking* emerge come fattore cruciale per l'evoluzione dei sistemi EPR tessili. L'integrazione di tecnologie digitali e di Intelligenza Artificiale – come previsto anche dalle più recenti disposizioni dell'AI Act – rappresenta un'opportunità per rafforzare la raccolta dei dati, la trasparenza dei flussi informativi e l'efficacia dei processi decisio-

[33] M. RENNA, *Diritto dell'ambiente*, il Mulino, Bologna, 2019, 145 ss.

[34] A. PAJNO, *La digitalizzazione del procedimento amministrativo*, in «Jus Publicum», 2023.

[35] C. ALESSI, *Amministrazione pubblica e innovazione tecnologica: profili giuridici*, in «Rivista Italiana di Informatica e Diritto», 2022.

nali, ma pone al contempo nuove sfide, legate all'armonizzazione dei framework giuridici, alla tutela dei dati e alla prevenzione della frammentazione del mercato unico europeo.

L'esperienza dell'EPR tessile mette dunque in luce una tensione strutturale tra efficienza e garanzie. Da un lato, l'uso di sistemi algoritmici appare indispensabile per assicurare l'effettività delle politiche ambientali in un settore ad alta complessità; dall'altro, tale uso rischia di comprimere le garanzie tradizionali del procedimento amministrativo, soprattutto in termini di trasparenza, motivazione e imputazione della decisione. In questo senso, l'EPR tessile si configura come un vero e proprio *stress test* per la capacità del diritto amministrativo di governare l'innovazione tecnologica senza rinunciare alle proprie categorie fondamentali[36].

Sotto questo profilo, il coordinamento tra disciplina ambientale e regolazione dell'Intelligenza Artificiale assume un rilievo decisivo. Le previsioni del Regolamento (UE) 2024/1689 in materia di governance, tracciabilità e responsabilità dei sistemi di IA offrono strumenti utili per rafforzare il controllo pubblico sull'uso degli algoritmi anche nell'attuazione dei regimi EPR. Tuttavia, l'effettività di tali strumenti dipende dalla loro integrazione con i principi del diritto amministrativo, chiamati a garantire che l'innovazione tecnologica si traduca in un rafforzamento, e non in un indebolimento, della legalità dell'azione amministrativa anche nei modelli decisionali algoritmicamente mediati.

6. Conclusioni

L'analisi svolta ha messo in evidenza come l'impiego dell'Intelligenza Artificiale nell'attuazione delle politiche ambien-

[36] S. CASSESE, cit.

tali, e in particolare nei sistemi di Responsabilità Estesa del Produttore applicati al settore tessile, incida in profondità sulle categorie fondamentali del diritto amministrativo, sollecitando una riflessione critica sulla trasformazione della funzione pubblica e sulla capacità dell'ordinamento di governare tali mutamenti senza rinunciare alle proprie garanzie strutturali.

Dall'esame del procedimento amministrativo emerge come l'algoritmo determini una progressiva ibridazione tra valutazione giuridica e razionalità tecnica, assumendo un ruolo sempre più rilevante nella selezione degli elementi decisionali.

L'esperienza della Responsabilità Estesa del Produttore (EPR) nel settore tessile si configura simultaneamente come uno *stress test* sistemico per il diritto amministrativo contemporaneo e come un laboratorio privilegiato di sperimentazione per l'amministrazione algoritmica. In questo ambito, caratterizzato da un'elevata complessità tecnica, decisioni seriali e una forte integrazione tra pubblico e privato, l'uso strutturale dell'Intelligenza Artificiale rende evidente la tensione tra le esigenze di *datafication* dell'azione pubblica e la salvaguardia delle garanzie tradizionali di legalità, responsabilità e controllo.

Questa dinamica trova un solido fondamento empirico nel modello *Eco-Circular Denim*[37], il quale dimostra come un quadro *smart* e *municipality-centred*, basato sull'integrazione tra eco-design e tracciabilità digitale, sia in grado di rendere operativa l'economia circolare nei contesti urbani, allineando le infrastrutture locali alle ambizioni delle politiche europee[38]. Tale sperimentazione trasforma l'algoritmo da semplice supporto tecnico a una vera e propria infrastruttura di potere, capace di orientare l'allocazione delle risorse e la distribuzione degli oneri. Di conseguenza, il caso del tes-

[37] C. PARENTE *et al.*, cit.

[38] D. PALAZZO, cit.

sile anticipa una metamorfosi necessaria per l'intero settore pubblico.

Tuttavia, l'emergere di queste logiche evidenzia la necessità di ripensare il ruolo dell'amministrazione: essa non deve essere solo il decisore finale, ma il soggetto responsabile del governo complessivo di processi decisionali tecnologicamente mediati, anticipando dinamiche destinate a estendersi a ogni altro ambito dell'azione pubblica[39].

Le criticità connesse alla trasparenza e alla motivazione dell'azione amministrativa algoritmica impongono un approccio sostanziale ai principi di legalità. La trasparenza, in particolare, deve evolvere da mera accessibilità del codice a strumento di controllo effettivo del potere. In questo senso, il Passaporto Digitale dei Prodotti (DPP) funge da dispositivo d'elezione: esso materializza la trasparenza richiesta dall'ordinamento garantendo dati strutturati e verificabili sulla composizione dei materiali, indispensabili per definire la responsabilità del produttore.

Particolarmente delicato si rivela, inoltre, il tema dell'imputazione della decisione e della responsabilità, già emerso nell'analisi dell'amministrazione algoritmica. In tali ipotesi, il rischio è quello di una responsabilità "diluita", nella quale il potere decisionale si frammenta lungo una filiera tecnico-organizzativa difficilmente riconducibile a un centro unitario di imputazione, rendendo indispensabile un rafforzamento dei presidi di governo pubblico dell'algoritmo, sia nella fase di progettazione sia in quella di utilizzo. In tale contesto, la responsabilità dell'amministrazione rischia di rimanere formalmente affermata ma sostanzialmente indebolita, nella misura in cui la decisione, pur imputata all'autorità pubblica, è in larga parte determinata da processi algoritmici progettati e gestiti al di fuori del circuito decisionale pubblico.

[39] F. PASQUALE, cit.

In questo contesto, il Regolamento (UE) 2024/1689 sull'Intelligenza Artificiale costituisce un passaggio di rilievo, nella misura in cui introduce obblighi di governance, tracciabilità e accountability idonei a incidere anche sull'uso dell'IA nell'azione amministrativa. Tuttavia, come l'analisi ha mostrato, l'AI Act non può essere considerato una risposta autosufficiente alle questioni poste dall'amministrazione algoritmica: la sua effettività dipende dalla capacità di integrazione con i principi e gli istituti del diritto amministrativo, chiamati a garantire che l'innovazione tecnologica si traduca in un rafforzamento – e non in un indebolimento – delle garanzie dell'ordinamento.

La sfida dell'amministrazione algoritmica risiede nella capacità di riaffermare la funzione delle categorie tradizionali in un ecosistema mutato. Trasparenza, motivazione e controllo giurisdizionale restano i pilastri irrinunciabili della legalità, anche quando il potere si esercita tramite software.

Il caso dell'EPR mostra come l'obiettivo sia garantire che l'innovazione tecnologica sia sempre accompagnata da una responsabilità politica e amministrativa certa, mantenendo integro il rapporto tra decisione pubblica e tutela degli interessi coinvolti. In definitiva, l'esperienza dell'EPR tessile mostra come il dato assuma oggi la natura di un asset strategico dell'azione amministrativa, non riducibile a mero input informativo. La sua elaborazione algoritmica incide sulla formazione della decisione pubblica, sulla distribuzione degli oneri economici e sull'orientamento dei comportamenti produttivi, configurandosi come infrastruttura giuridica della governance ambientale. In tale prospettiva, la questione centrale non riguarda soltanto l'efficienza dei sistemi *data-driven*, ma la capacità dell'ordinamento di governare il dato come luogo di esercizio del potere pubblico, preservando trasparenza, imputabilità e controllo. Il dato diviene così il nuovo asset strategico sul quale si

misura la tenuta delle garanzie amministrative nell'esercizio del potere pubblico nell'ecosistema digitale.

L'intelligenza artificiale e il diritto: AI ruled ed AI ruler

di *Anna Pirozzoli*[1]

SOMMARIO: 1. L'AI ruled – 2. L'influenza dell'intelligenza artificiale sul diritto: l'AI ruler – 3. E se l'intelligenza artificiale irrompesse nei giudizi della Corte costituzionale? – 4. Conclusioni.

1. L'AI ruled

Occuparsi dell'intelligenza artificiale (AI) rappresenta un mandato irrinunciabile ormai anche per i giuristi che devono misurarsi con l'intensità con cui i sistemi di AI possono influenzare i diritti, non solo nella loro declinazione di diritti soggettivi bensì, più in generale, il diritto con riferimento non marginalmente anche alle modalità di svolgimento dei giudizi[2].

Nell'affrontare questo tema è opportuno assumere un atteggiamento cauto per quanto non troppo timoroso, che riconosca l'inarrestabilità dei processi di innovazione tecnologica e la loro capacità di influenzare i meccanismi sociali in maniera sempre più penetrante[3]. Occorre osservarne i cambiamenti, e ove possibile

[1] Professore Ordinario di Istituzioni di diritto pubblico, Università degli Studi di Niccolò Cusano.

[2] T. GROPPI, *Alle frontiere dello Stato costituzionale: innovazione tecnologica e intelligenza artificiale,* in *Consulta online,* III/2020, pp. 675; M. LUCIANI, *La sfida dell'intelligenza artificiale,* in «Rivista AIC», 12, 2023, 3–4; V.A. LAPTEV, D.R. FEYZRAKHMANOVA, *Application of Artificial Intelligence in Justice: Current Trends and Future Prospects*, in «HumCentric Intelligence System», 2024, 1–12.

[3] Sono trascorsi quasi settant'anni dal momento in l'intelligenza artificiale fece il suo ingresso in un contesto scientifico tramite la proposta di ricerca presenta-

prevederli, per predisporre strumenti di regolamentazione capaci di contenere quelle forme di intelligenza artificiale che rischiano di essere dominanti rispetto all'intelligenza umana stessa[4], pur nella prospettiva di un'inevitabile convivenza – per alcuna dottrina un vero e proprio "intreccio"[5] – dell'uomo con le tecnologie emergenti[6].

L'ampia porzione di sconosciuto riguardo l'AI suggerisce per il momento di prediligere, giudiziosamente, regolamentazioni costituite da principi generali muovendo dai quali occorre operare in un'ottica di construption[7], di accoglimento e accettazione delle

ta da J. McCarthy e co. per la Conferenza di Dartmouth: J. MCCARTHY, M.L. MINSKY, N. ROCHESTER, C.E. SHANNON, *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*, Dartmouth College, Hanover, New Hampshire, 1955, il testo è disponibile al seguente indirizzo *jmc.stanford.edu.*

[4] U. RUFFOLO, A. AMIDEI, I*ntelligenza artificiale e diritti della persona: le frontiere del "transumanesimo",* in *Giur. It.,* 7/2019, pp. 1657; I. BREMMER, M. SULEYMAN, *The AI Power Paradox. Can States Learn to Govern Artificial Intelligence - Before It's Too Late?*, in «Foreign Affairs»*,* Sept/Oct 2023: «Farther out on the horizon lurks the promise of artificial general intelligence, the still uncertain point where AI exceeds human performance at any given task, and the (admittedly speculative) peril that AGI could become self-directed, self-replicating, and self-improving beyond human control. All these dangers need to be factored into governance architecture from the outset».

[5] Di un *entanglement* tra AI, esseri umani, società e natura parlano A. KÖVES, K. FEHER, L. VICSEK *et al.*, *Entangled AI: artificial intelligence that serves the future*, in «AI & Soc», Aug. 2024, 1–12.

[6] B. BRAUNSCHWEIG, M. GHALLAB, *Reflections on Artificial Intelligence for Humanity. Lecture Notes*, Springer, Cham, 2021; A. DENGEL, O. ETZIONI, N. DECARIO, H. H. HOOS, L. FEI-FEI, J. TSUJII, P. TRAVERSO, *Next Big Challenges in Core AI Technology. Reflections on Artificial Intelligence for Humanity*, Springer, Cham, 2021, pp. 90-115.

[7] M. LUCIANI, *Può il diritto disciplinare l'intelligenza artificiale? Una conversazione preliminare*, in «Bilancio Comunità Persona», 2, 2023, «Si potrebbero dunque introdurre almeno alcuni princìpi generali, sebbene l'estrema varietà del fenomeno possa astrattamente suggerire l'adozione di discipline specifiche per i vari campi di applicazione dell'intelligenza artificiale. In realtà, la presenza di

tecnologie emergenti il cui rallentamento non è previsto né tantomeno auspicabile[8].

Tuttavia, i termini di quest'opera di consolidamento giuridico coinvolgono il diritto e l'intelligenza artificiale in modi diversi in funzione della prospettiva da cui vengono osservati.

Per un verso c'è l'aspetto degli strumenti giuridici, delle modalità e dell'intensità di regolamentazione dell'intelligenza artificiale, considerata qui come oggetto di disciplina, dunque potenzialmente ruled, tramite hard law o soft law a seconda dell'intento di creare maglie legislative più o meno ampie entro cui contenere questo nuovo fenomeno tecnologico ed i suoi effetti[9].

Al riguardo, il principale riferimento normativo in tema di disciplina dell'intelligenza artificiale è al momento senz'altro il Regolamento UE 2024/1689 del Parlamento Europeo e del Consiglio, noto come AI Act, in vigore dal 1° agosto 2024 benché pienamente applicabile nel 2026. È considerato il primo legal framework che delinea uno standard globale riguardante la produzione, la distribuzione e le forme di utilizzo dei sistemi di AI di cui vengono presi in considerazione i rischi tramite una scala di graduazione che distingue obblighi o divieti in funzione del minimal, limited, high o unacceptable risk[10].

princìpi generali faciliterebbe la regolazione delle future novità, da attendersi profonde e in rapida successione, che una regolazione analitica sarebbe costretta (sempre in ritardo) a inseguire».

[8] M.U. SCHERER, *Regulating Artificial Intelligence Systems: Risks, Challenges, Com-petencies, and Strategies*, in «Harvard Journal of Law and Technology», 29, 2, 2016, 373.

[9] O. POLLICINO, *Stato di diritto, diritto costituzionale e intelligenza artificiale*, in «LeCostituzionaliste.it», 20 giugno 2024.

[10] C. NOVELLI, F. CASOLARI, A. ROTOLO, M. TADDEO, L. FLORIDI, *AI Risk Assessment: A Scenario-Based, Proportional Methodology for the AI* Act, in «DISO», 2024, 11; J. CHAMBERLAIN, *The Risk-Based Approach of the European Union's Proposed Artificial Intelligence Regulation: Some Comments from a Tort Law*

La "combinazione della probabilità del verificarsi di un danno e la gravità del danno stesso" – questa è la definizione di rischio nell'art. 3 dell'AI Act – rappresenta l'elemento su cui si fonda l'insieme degli obblighi e dei divieti relativi ai sistemi di AI, e suggerisce la volontà normativa di un lungimirante processo di regolamentazione, potenzialmente in grado di operare in prospettiva futura o, forse più correttamente, futuribile.

Peraltro, appare apprezzabile la scelta di istituire allo stesso tempo un AI Pact che incoraggia le organizzazioni del settore ad anticipare e avviare prima della scadenza legale l'attuazione del dettato normativo del Regolamento UE 2024/1689. Questo approccio, oltre alla decisione di creare ambienti applicativi di prova costituiti dalle regulatory sandbox, garantisce l'istituzione di un'area di sperimentazione normativa a livello nazionale per agevolare lo sviluppo e le prove di sistemi di intelligenza artificiale innovativi prima che tali sistemi siano immessi sul mercato o altrimenti messi in servizio.

Si tratta, insomma, di una forma di "legislazione future-proof"[11], che consente alle norme di adattarsi ai cambiamenti tec-

Perspective, in «European Journal of Risk Regulation», 14, 2023, 1–13. La necessità di un intervento normativo urgente è evidenziata anche dall'*Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence* firmato il 30 ottobre 2023 dal Presidente Biden il quale afferma: «My Administration places the highest urgency on governing the development and use of AI safely and responsibly, and is therefore advancing a coordinated, Federal Government-wide approach to doing so. The rapid speed at which AI capabilities are advancing compels the United States to lead in this moment for the sake of our security, economy, and society». Il testo completo è disponibile all'indirizzo *www.whitehouse.gov.*

[11] S. LARSSON, J. HILDÉN, K. SÖDERLUND, *Between Regulatory Fixity and Flexibility in the EU AI Act*, Lund University, Lund, 2024, 17; G. BAS, C. SALINAS, R. TINOCO, J. SEVILLA, *The EU AI Act: A pioneering effort to regulate frontier AI?* in «Inteligencia Artificial», 2024, 55–64; G. ROBINSON, S. DE VRIES, B. DUIV-

nologici, orientando gli Stati membri nella direzione di un approccio metodologico cd. foresight, teso alla previsione e all'anticipazione di scenari futuri[12].

2. L'influenza dell'intelligenza artificiale sul diritto: l'AI ruler

Nella seconda prospettiva di analisi, quella relativa alla capacità dell'AI di incarnare il ruolo di ruler, è necessario distinguere i modi con cui i sistemi di intelligenza artificiale sono in grado di influenzare il diritto, ad esempio condizionando i processi di amministrazione della giustizia, orientando i giudici nella stesura delle sentenze, e finanche il giudice costituzionale nell'opera di interpretazione e bilanciamento degli interessi in gioco[13].

A riguardo già l'AI Act ha evidenziato la particolare pericolosità dei sistemi di intelligenza artificiale in questo settore, includendoli tra quelli ad "alto rischio", in considerazione del loro impatto potenzialmente significativo sulla democrazia, sullo Stato di diritto,

ENVOORDE, *Future-proof Regulation and Enforcement for the Digitalised Age*, in «European Papers», 8, 2023, 579–582.

[12] Il *foresight* e l'*adaptability* sono considerati due elementi irrinunciabili del *future-proof approach*, a riguardo si veda S. RANCHORDÁS, M. VAN'T SCHIP, *Future-proofing legislation for the digital age*, in S. RANCHORDÁS, Y ROZNAI, *Time, Law, and Change. An Interdisciplinary Study*, Hart Publishing, Oxford, 2020. Sulla necessità di una normativa sovranazionale costruita con un approccio maggiormente orientato al *foresight* cfr. G. NATALE, A.M. LAZZÈ, *AI Act, il Regolamento europeo sull'intelligenza artificiale: punti di forza e punti di debolezza*, in «Rassegna Avvocatura dello Stato», 3, 2023, 139, secondo gli autori «tramite la sua interdisciplinarità, risulta a maggior ragione una strada più che valida da seguire nel tentativo di regolamentare adeguatamente il fenomeno in questione già da oggi, nell'ottica di ciò che sarà».

[13] Su cui si veda L. DI MAJO, *Le incognite sull'utilizzo dell'Intelligenza Artificiale nel processo costituzionale in via incidentale*, in «La Rivista Gruppo di Pisa», 2, 2023, 17.

sulle libertà individuali e sul diritto a un ricorso effettivo e a un giudice imparziale.

Tra i sistemi che maggiormente preoccupano i giuristi ci sono indubbiamente i sistemi di cd. *predictive justice*, per i quali sembra più corretto parlare di attività di previsione statistica fondate su calcoli probabilistici, e non di predittività, termine che sembra voglia richiamare una sorta di chiaroveggenza che evidentemente non è propria di un algoritmo[14].

La giustizia predittiva può consentire di selezionare i metadati disponibili confezionando – induttivamente – un argomento esperienziale su cui fondare la decisione; o per via deduttiva simulare il ragionamento fondato sul sillogismo giudiziario descritto dall'art. 12 delle Preleggi del Cod. civ.[15], giungendo a concretizzare quell'ipotesi di applicazione alla giustizia dei modelli matematici presagita già nella metà del 1600 da Gottfried Wilhelm Leibniz[16].

Se per un verso questa "opportunità tecnologica" lascia prospettare alcuni vantaggi nella direzione di una riduzione dei tempi dell'amministrazione della giustizia – pur senza trascurarne i rischi[17] – nell'aumento del ricorso dei sistemi conciliativi, o anche,

[14] Su cui v. G. CARLOTTI, *La giustizia predittiva e le fragole con la panna,* in «Giustizia-amministrativa.it», 2023.

[15] Cfr. V. TENORE, *I robot in giudizio! Considerazioni sull'utilizzo di intelligenza artificiale (AI) da parte del magistrato: abdicare al ragionamento e alla riflessione quotidiana a favore di un robot togato ovvero un "cretino digitale", è un bene per la giustizia e per i suoi attori e destinatari?* in «Rivista della Corte dei conti», 4, 2023, 3.

[16] G.W. LEIBNIZ, *Dissertatio de arte combinatoria, in qua, ex arithmeticae fundamentis*, Lipsiae, 1666. Su cui A. TRIGGIANO, *Il "sogno" di una giustizia predittiva tra Seicento e Ottocento*, in «Rivista italiana per le scienze giuridiche», 13, 2022, 105. Sulle relazioni tra giustizia predittiva e utilizzo dei modelli matematici v. L. VIOLA, *Overruling e giustizia predittiva,* Giuffrè, Milano, 2020, 109.

[17] Sui rischi derivanti dalla rapidità nel diritto e nell'amministrazione della giustizia si veda M. LUCIANI, *La decisione giudiziaria robotica*, in «Rivista AIC», 3,

secondo alcuni, in un'aspirazione (discutibile) di oggettività nella composizione decisoria[18]; peraltro occorre rilevare il pericolo di un'assuefazione del giudice alla previsione algoritmica, tale da sopire quell'opera di intelletto necessaria a formulare una decisione umanamente ponderata.

Anche se solo si prova ad assegnare alla tecnologia il ruolo di "oggetto giudicante" quale mero supporto al giudice-uomo, possiamo prospettare taluni scenari non molto rassicuranti per gli effetti sui giudizi e sul ruolo del giudice[19].

Senza spingere l'immaginazione a figurare sistemi di intelligenza artificiale troppo evoluti e ancora ignoti, già solo la robotizzazione derivante dall'utilizzo dei repertori giurisprudenziali si mostra gravida di dubbi e interrogativi, che invece di confortare nella direzione di una incontrovertibilità del giudizio, evidenziano i numerosi limiti derivanti dal loro utilizzo.

Primo tra tutti emerge il problema del popolamento delle informazioni nei repertori: le modalità di inserimento dei dati giurisprudenziali, la forchetta temporale adottata, la difficoltà nel definire un criterio di risoluzione dei contrasti giurisprudenziali, ed il vincolo alla giurisprudenza precedente in violazione dell'art. 101 Cost., evidenziano i limiti di una decisione giudiziaria robotica che, allo stato dell'arte, sembra mostrare più svantaggi che vantaggi[20].

2018, 874.

[18] D. KAHNEMAN, O. SIBONY, C.R. SUNSTEIN, *Noise, A flaw in human judgment*, William Collins, London, 2022.

[19] Si contano numerose sperimentazioni di sistemi di intelligenza artificiale applicati alla giustizia, tra cui *Lider Lab* presso la Scuola superiore "Sant'Anna" di Pisa, le esperienze pilota presso la Corte di appello di Brescia, il Tribunale di Genova, la Corte d'appello di Venezia, il Tribunale di Firenze, l'accordo di collaborazione tra il CED della Corte di Cassazione e la Scuola universitaria superiore di Pavia (IUSS), su cui G. CARLOTTI, *op. cit.,* 1.

[20] Sulla "decisione giudiziaria robotica" si veda M. LUCIANI, *La decisione giudiziaria robotica*, cit., pp. 886, secondo l'A. «è ragionevole pensare che mettere in

Inoltre, la riduzione del lavoro umano nell'atto decisionale alimenterebbe a sua volta un adattamento progressivo del giudice alla macchina, consentendogli di adagiarsi più o meno inconsapevolmente al suggerimento del giudizio robotico (forse già un "giudizio anticipato"), riducendo gradualmente l'autonomia e l'indipendenza del giudice, invece garantite costituzionalmente[21].

Per di più questo è un rischio che non riguarda solo l'organo giudicante, ma finisce per travolgere anche il magistero dell'avvocato che vede una forma d'interferenza dell'AI quantomeno sulle modalità di svolgimento del contraddittorio, oltre che sul ruolo del difensore potenzialmente indebolito dai dominanti automatismi dei sistemi di intelligenza artificiale, fino ad insidiare i principi generali del giusto processo richiamati in Costituzione all'art. 111[22].

Ma vi è di più, poiché oltre ai principi giuridici – su cui appare a questo punto innegabile l'ampio rischio di un'incidenza dei sistemi di intelligenza artificiale – occorre considerare anche l'im-

pratica l'idea estrema del robot-giudice abbia più svantaggi che vantaggi. Come ben sanno i matematici che si occupano del calcolo delle probabilità, la previsione non è predizione, sicché quando chiediamo che le decisioni giudiziarie siano prevedibili eleviamo una pretesa di certezza giuridica, non di infallibilità matematica (ammesso - e nient'affatto concesso - che il robot possa essere infallibile)».

[21] G. DE MINICO, *Giustizia e intelligenza artificiale: un equilibrio mutevole*, in *Rivista AIC,* 2/2024, pp. 98. Su autonomia e indipendenza dei giudici cfr. almeno S. BARTOLE, *Autonomia e indipendenza dell'ordine giudiziario*, Cedam, Padova, 1964, 4; G. VOLPE, *Consiglio superiore della magistratura*, in *Enc. Dir.,* Agg. IV, Giuffrè, Milano, 2000, 380; N. ZANON, F. BIONDI, *Il sistema costituzionale della Magistratura*, Zanichelli, Bologna, V ed., 2019, 39. Sulla distinzione tra le tecnologie che assistono la decisione e quelle che sostituiscono la decisione si veda A. SIMONCINI, *Il linguaggio dell'intelligenza artificiale e la tutela costituzionale dei diritti*, in «Rivista AIC», 2, 2023, 13 ss.

[22] Di un *cyber-lawyer* parla L. DI MAJO, *"Mi rimetto alla clemenza dell'intelligenza artificiale". L'AI-jurisdicional process e il cyber-lawyer*, in «Giustizia insieme», 24 febbraio 2024.

patto sui principi etici che integrano i primi aggiungendovi una differente prospettiva di responsabilità[23].

Sono cinque, in particolare, quelli enunciati nella "Carta etica sull'utilizzo dell'intelligenza artificiale nei sistemi giudiziari e negli ambiti connessi" adottata dall'European Commission for the efficiency of justice (CEPEJ) nel corso della 31ª Riunione plenaria (Strasburgo, 3-4 dicembre 2018)[24]: il principio del rispetto dei diritti fondamentali; il principio di non-discriminazione; il principio di qualità e sicurezza; il principio di trasparenza, imparzialità ed equità; ed il principio del "controllo da parte dell'utilizzatore".

L'utilizzo etico degli strumenti di AI nei sistemi giudiziari impone quindi il rispetto di tali principi, simbioticamente all'osservanza dei principi giuridici poiché entrambi vòlti ad assicurare, in particolare, la compatibilità dell'elaborazione di strumenti e servizi di AI con i diritti fondamentali garantiti dalla Convenzione europea sui diritti dell'uomo (CEDU); prevenire lo sviluppo di qualsiasi discriminazione tra persone o gruppi di persone; e, non da ultimo, raggiungere un equilibrio tra la proprietà intellettuale e l'esigenza di trasparenza, imparzialità, equità e integrità intellettuale. Quest'ultimo, peraltro, riguarda particolarmente le ipotesi di strumenti che possono avere conseguenze giuridiche o che possono condizionare significativamente la vita delle persone, aspetto che il Consiglio di Stato ha voluto corroborare operando un bilan-

[23] U. RUFFOLO (a cura di), *Intelligenza artificiale. Il diritto, i diritti, l'etica*, Giuffrè, Milano, 2020.

[24] Il testo della "Carta etica sull'utilizzo dell'intelligenza artificiale nei sistemi giudiziari e negli ambiti connessi" è disponibile all'indirizzo *www.coe.int/en/web/cepej.* Sulla "Carta etica sull'utilizzo dell'intelligenza artificiale nei sistemi giudiziari e negli ambiti connessi" cfr. F. CAMILLIERI, *Gli algoritmi predittivi alla luce dei Principi delineati nella European Ethical Charter*, in «Quaderno monografico abbinato al fascicolo 2021/2 de La Rivista Gruppo di Pisa», 313.

ciamento tra la tutela della proprietà intellettuale dell'algoritmo e il principio di trasparenza in favore del principio di trasparenza algoritmica secondo il quale "il meccanismo attraverso il quale si concretizza la decisione robotizzata (ovvero l'algoritmo) debba essere "conoscibile", secondo una declinazione rafforzata del principio di trasparenza, che implica anche quello della piena conoscibilità di una regola espressa in un linguaggio differente da quello giuridico.

Tale conoscibilità dell'algoritmo deve essere garantita in tutti gli aspetti: dai suoi autori al procedimento usato per la sua elaborazione, al meccanismo di decisione, comprensivo delle priorità assegnate nella procedura valutativa e decisionale e dei dati selezionati come rilevanti"[25].

3. E se l'intelligenza artificiale irrompesse nei giudizi della Corte costituzionale?

Se c'è un giudizio in cui l'opera di bilanciamento degli interessi non può essere demandata all'intelligenza artificiale è il giudizio assegnato alla Corte costituzionale in cui il libero convincimento dei giudici costituzionali si costituisce al termine di un meticoloso esercizio di interpretazione, ove l'elemento valutativo assiologico rappresenta «parte integrante del ragionamento giuridico "scientifico"»[26], nel cui processo ermeneutico dunque – secondo la teoria

[25] In tal senso Cons. Stato, sez. VI, 8 aprile 2019, n. 2270, par. 13.1., su cui A. FORMISANO, *L'impatto dell'intelligenza artificiale in ambito giudiziario sui diritti fondamentali,* in «Federalismi.it», 22, 2024, 133. Sul principio di trasparenza come "pietra miliare" nella delineazione delle regole cui deve conformarsi l'applicazione dei sistemi algoritmici cfr. G. LO SAPIO, *La trasparenza sul banco di prova dei modelli algoritmici*, in «Federalismi.it», 11, 2021, 239.

[26] G. RAZZANO, *Riflessioni su Emilio Betti filosofo e costituzionalista, suo malgrado*, in «Nomos», 1, 2017, 13.

di Emilio Betti – si uniscono inscindibilmente il momento logico, quello teleologico ed, appunto, quello assiologico[27].

Il canone della ragionevolezza, utilizzato come «complemento e in appoggio a qualunque altro principio costituzionale richiamato a parametro del giudizio della Corte»[28], più di ogni altro richiede una "funzione attiva" della mente umana che non è replicabile dall'opera della macchina, particolarmente quando emerge la necessità di elaborare "giudizi di valore" a discapito di un'opera esclusivamente logico-deduttiva.

Ed infatti appare del tutto improbabile consegnare ad un algoritmo un giudizio di ragionevolezza che la stessa Corte costituzionale, nella sentenza n. 1130 del 1988, ha chiarito debba svolgersi attraverso "ponderazioni relative alla proporzionalità dei mezzi prescelti dal legislatore nella sua insindacabile discrezionalità rispetto alle esigenze obiettive da soddisfare o alle finalità che inten-

[27] E. BETTI, *Teoria generale della interpretazione*, (1955) edizione corretta e ampliata a cura di G. CRIFÒ, vol. II, Giuffrè, Milano, 1990, 802. Sull'interpretazione costituzionale A. PACE, *Interpretazione costituzionale e interpretazione per valori,* in G. AZZARITI (a cura di), *Interpretazione costituzionale*, Giappichelli, Torino, 2007, 83; F. MODUGNO, *Interpretazione per valori e interpretazione costituzionale*, in G. AZZARITI (a cura di), *Interpretazione costituzionale*, 51; G. U. RESCIGNO, *Interpretazione costituzionale e positivismo giuridico*, in G. AZZARITI (a cura di), *Interpretazione costituzionale*, op. cit., 21; F. MODUGNO, *Sulla specificità dell'interpretazione costituzionale*, in ID., *Scritti sull'interpretazione costituzionale*, ESI, Napoli, 2008, 258.

[28] M. CARTABIA, *I principi di ragionevolezza e proporzionalità nella giurisprudenza costituzionale italiana*, intervento presentato all'incontro trilaterale tra la Corte costituzionale italiana, la Corte costituzionale spagnola e il Tribunale costituzionale portoghese, Palazzo della Consulta 24–26 ottobre 2013, il testo è disponibile all'indirizzo *www.cortecostituzionale.it.* Sulla pervasività del canone della ragionevolezza quale principio onnipresente nella giurisprudenza costituzionale si veda L. PALADIN, *Ragionevolezza (principio di)*, in *Enc. Dir.*, Aggiornamento, I, Giuffrè, Milano, 1997, 899. Sulla Corte costituzionale come custode della ragionevolezza cfr. A. MORRONE, *Il custode della ragionevolezza*, Giuffrè, Milano, 2001.

de perseguire, tenuto conto delle circostanze e delle limitazioni concretamente sussistenti".

La peculiarità del caso concreto, dunque, è questo il nucleo intorno al quale ruota la legittimazione del potere dei giudici costituzionali di interpretare – secondo ragionevolezza – senza cedere agli "automatismi legislativi"[29], soprattutto quando occorre operare una scelta etica che impone l'uso dell'intelligenza sociale del giudice[30], quella che permette di considerare il contesto, evitando peraltro forme di discriminazione che probabilmente un algoritmo non saprebbe riconoscere e controllare correttamente.

L'utilizzo di algoritmi predittivi in fase decisionale, infatti, unitamente ad una ridotta "sorveglianza umana" derivante dall'adattamento ad una valutazione tecnologicamente orientata, potrebbe generare nuove forme di discriminazione derivanti dalla "qualità" della programmazione algoritmica e dalla preferenza del peso assegnato ai dati/valori inseriti nel sistema (fondati, ad esempio, su modelli cd. garbage in garbage out), con il conseguente condizionamento nella formulazione dei risultati[31].

Del resto, non si può rischiare di esporre anche il processo costituzionale ad un così elevato rischio di pregiudizi (bias) algorit-

[29] M. CARTABIA, *op. ult. cit.,* pp. 12. In argomento A. SIMONCINI, S. SUWEIS, *Il cambio di paradigma nell'intelligenza artificiale e il suo impatto sul diritto costituzionale*, in «Rivista di filosofia del diritto», 1, 2019, 102; L. PACE, *L'adeguatezza della legge e gli automatismi. Il Giudice delle leggi fra norma "astratta" e caso "concreto"*, Editoriale Scientifica, Napoli, 2020.

[30] Di intelligenza sociale del giudice parla F. SUCAMELI, *Una proposta di discussione,* in «Diritto & Conti - "Bilancio Comunità Persona"», 2, 2023, 20. Sulle questioni etiche sottoposte alla Corte costituzionale v. L. DI MAJO, *Le incognite sull'utilizzo dell'Intelligenza Artificiale nel processo costituzionale in via incidentale,* cit. 41.

[31] Su cui A. SIMONCINI, *L'algoritmo incostituzionale: intelligenza artificiale e il futuro delle libertà*, in «BioLaw Journal», 1, 2019, 84; M. D'AMICO, *Linguaggio discriminatorio e garanzie costituzionali*, in «Rivista AIC», 1, 2023, 243.

mici già evidenziato in Italia ad esempio in alcuni procedimenti amministrativi[32] e dal Tribunale di Bologna nel caso Deliveroo[33]; oltre che in altri sistemi giudiziari[34], ad esempio dalla Corte distrettuale dell'Aja sulla tecnica SIRI utilizzata dal governo dei Paesi Bassi[35], ma anche dalla Corte d'Appello dell'Inghilterra e del Galles sull'impatto dei bias prodotti dal software di riconoscimento

[32] Cfr. ad esempio TAR Lazio, sez. III bis, 22 marzo 2017, n. 3769; TAR Lazio, sez. III bis, 21 marzo 2017, n. 3742; TAR Lazio, sez. III bis, 10 settembre 2018, nn. 9224-9930; TAR Lazio, sez. VI, 13 dicembre 2019, nn. 8472- 8474; Consiglio di Stato, sez. VI, 8 aprile 2019, n. 2270; Cons. di Stato, sez. VI, 4 febbraio 2020, n. 881; TAR Lazio, sez. III bis, 24 giugno 2021, n. 7589. In argomento G. AVANZINI, *Decisioni amministrative e algoritmi informatici, Predeterminazione, analisi predittiva e nuove forme di intellegibilità*, Editoriale Scientifica, Napoli, 2019; A. SIMONCINI, *op. ult. cit.,* pp. 73-76; G. PESCE, *Il Consiglio di Stato ed il vizio della opacità dell'algoritmo tra diritto interno e diritto sovranazionale*, in «Giustizia-amministrativa.it», 2020; P. FALLETTA, *Alla ricerca di un paradigma giuridico della decisione amministrativa algoritmica*, in «Amministrazione in Cammino», 31 maggio 2022.

[33] Su cui D. TESTA, *La discriminazione degli algoritmi: il caso Deliveroo, Trib. Bologna, 31 dicembre 2020*, in «Ius in Itinere», 26 gennaio 2021.

[34] Su cui v. T. GROPPI, *op. cit.,* 677; E. STRADELLA, *Stereotipi e discriminazioni: dall'intelligenza umana all'intelligenza artificiale*, in AA.VV., *Liber Amicorum per Pasquale Costanzo*, disponibile su «ConsultaOnline», 30 marzo 2020; M. D'AMICO, *Una parità ambigua. Costituzione e diritti delle donne*, Raffaello Cortina, Milano, 2020, pp. 313; A. D'ALOIA, *Il diritto verso "il mondo nuovo". Le sfide dell'intelligenza artificiale,* in A. D'ALOIA (a cura di), *Intelligenza artificiale e diritto. Come regolare un mondo nuovo,* Franco Angeli, Milano, 2020, pp. 7. Per un'analisi accurata dei singoli casi si veda C. NARDOCCI, *Intelligenza artificiale e discriminazioni*, in «Rivista Gruppo di Pisa», 3, 2021, 54; J. CASILLAS, *Bias and Discrimination in Machine Decision-Making Systems*, in F. LARA, J. DECKERS, (eds), *Ethics of Artificial Intelligence. The International Library of Ethics, Law and Technology*, vol 41. Springer, Cham, 2023, 13.

[35] Sulla pronuncia della Corte distrettuale dell'Aja, n. C/09/550982/HAZA 18-388, 5 febbraio 2020 si veda A. RACHOVITSA, N. JOHANN, *The Human Rights Implications of the Use of AI in the Digital Welfare State: Lessons Learned from the Dutch SyRI Case*, in «Human Rights Law Review», 2022, 1.

facciale AFR Locate utilizzato dalle forze di polizia per ragioni di sorveglianza pubblica e di law enforcement[36].

Emblematico, in argomento, anche il caso di Eric L. Loomis condannato dal Tribunale circondariale di La Crosse anche sulla base delle elaborazioni del software di valutazione del rischio di recidiva e pericolosità sociale COMPAS, utilizzato legittimamente – secondo la Supreme Court del Wisconsin – poiché non considerato come l'elemento esclusivo di decisione, bensì quale supporto tecnico, non vincolante, nelle valutazioni del Tribunale[37].

Percorrendo questa direzione si fa strada l'idea dell'algoritmo come una sorta di "esperto" esterno, in grado di fornire elementi tecnico-scientifici e di acquisire informazioni attinenti a specifiche discipline. Tale prospettiva richiama la possibilità di consultare gli algoritmi anche nel processo costituzionale in veste di esperti di chiara fama – quelli previsti dopo la modifica delle Norme integrative per i giudizi davanti alla Corte costituzionale[38] – rispetto ai quali «non v'è ragione di non ritenere possibile l'utilizzo di un algoritmo non quale esperto, piuttosto attraverso l'esperto interpel-

[36] Corte d'Appello dell'Inghilterra e del Galles, *R (Bridges) vs CC South Wales & ors*, 11 agosto 2020, su cui K. NEUBURG, S. NESTLER, S. QUADFLIEG, *(Dis)Obedience in Digital Societies: Perspectives on the Power of Algorithms and Data*, Verlag, Berlin, 2022.

[37] Su cui S. CARRER, *Se l'amicus curiae è un algoritmo: il chiacchierato caso Loomis alla Corte Suprema del Wisconsin,* in «Giurisprudenza Penale Web», 4, 2019.

[38] Su cui A. IANNUZZI, *La Camera di consiglio aperta agli esperti nel processo costituzionale: un'innovazione importante in attesa della prassi*, in «Osservatorio costituzionale», 2, 2020, 13; S. FINOCCHIARO, *Verso una giustizia costituzionale più "aperta": la Consulta ammette le opinioni scritte degli "amici curiae" e l'audizione di esperti di chiara fama*, in «Sistema penale», 23 gennaio 2020; L. DI MAJO, *Gli esperti nei processi decisionali di legislatore e Corte costituzionale. Contributo allo studio del rapporto tra scienza, politica e diritto*, Jovene, Napoli, 2023, 103.

lato dal Collegio, comunque con conseguenze non irrilevanti sulla formazione del libero convincimento del Giudice relatore»[39].

4. Conclusioni

Se il test di Alan Turing ha mostrato l'incapacità della macchina di "gustare le fragole con la panna"[40], appare evidente quanto la tecnologia oramai stia imboccando una corsia di accelerazione che sembra suggerire una capacità sempre più ampia di intervenire nei processi decisionali, non da ultimo in quelli che riguardano l'amministrazione della giustizia. Ciononostante, non è auspicabile un rallentamento della tecnologia se si vuole onorare pienamente il dettato dell'art. 33 della Costituzione italiana[41], ma senz'altro occorre sorvegliarne le evoluzioni, consentendo alla ragione umana la possibilità di arginare le tracimazioni dell'intelligenza artificiale, soprattutto nella sua declinazione di AI ruler.

I tentativi di regolamentazione dell'AI – a livello nazionale e sovranazionale – contribuiscono solo parzialmente a delimitare la portata dei rischi di predominio della macchina sull'uomo soprattutto perché, nonostante ogni tentativo del legislatore di immaginare ciò che sarà, si dettano regole su qualcosa che non possiamo realmente prevedere, soprattutto non possiamo prevederne gli effetti sui sistemi giuridici e sui diritti (anche i "nuovi diritti") dell'uomo[42].

[39] L. DI MAJO, *Le incognite sull'utilizzo dell'Intelligenza Artificiale nel processo costituzionale in via incidentale*, cit., 21.

[40] A.M. TURING, *Computing machinery and intelligence*, in «Mind», 59, 1950, 433.

[41] F. CIRILLO, *Neuroprivacy. Problemi teorici e prospettive costituzionali,* Giappichelli, Torino, 2023, 126.

[42] L. RINALDI, *Intelligenza artificiale, diritti e doveri nella Costituzione italia-*

Probabilmente la soluzione alberga in una formula che potremmo definire diplomatica, ossia un'alleanza di cervelli, quello umano e quello tecnologico, possibile solo coltivando l'auspicio suggerito dal titolo di un libro di Jeremy Bernstein: non solo macchine intelligenti, ma soprattutto uomini intelligenti che sappiano interpretarle e governarle[43].

na, in «DPCE Online», 1, 2022, pp. 206.

[43] J. BERNSTEIN, *Uomini e macchine intelligenti*, G. LONGO (trad.it.), Adelphi, Milano, 1990.

Intelligenza Artificiale e Codice dei Contratti Pubblici

di *Francesca Salvatore**

SOMMARIO: 1. Premessa. – 2. Contributo umano e responsabilità dell'"operatore" privato e pubblico. – 3. L'attendibilità del modello. – 4. L'imputazione della responsabilità. – 5. Rischi e responsabilità. – 6. L'illusione dell'imparzialità computazionale nella pubblica amministrazione. – 6.1. L'illusione epistemologica: il dato come costruzione sociale. – 6.2. L'etica computazionale: la responsabilità tra scelta morale e calcolo automatico. – 6.3. Efficienza *versus* giustizia: antica dicotomia tra principi e regole applicabili. – 7. Attività amministrativa e algoritmi. – 8. Il nodo della trasparenza: l'opacità del "decision-making" automatizzato. – 9. Appalti Digitali, al via la Rivoluzione 2.0: AgID vara le nuove regole tecniche per le piattaforme. – 10. Esperienze applicative: tra sperimentazione e cautela. – 11. Sperimentazioni normative: l'hub dei contratti pubblici. – 12. Conclusioni: per una filosofia della responsabilità nell'era algoritmica.

1. Premessa

In tema di applicazione dell'intelligenza artificiale al mondo dei contratti pubblici viene spesso ricordato che l'art. 30 del nuovo Codice dei contratti, d.lgs. 36/2023, rappresenta una novità assoluta nello scenario normativo nazionale in quanto ha introdotto l'automazione delle attività amministrative nel ciclo di vita dei contratti pubblici: dalla programmazione sino alla fase di esecuzione. Contestualmente, il legislatore ha cercato di calibrare alcuni ele-

* Dottoranda in Società in mutamento: politiche, diritti e sicurezza, Università degli Studi della Tuscia.

menti in discussione in modo che il "funzionamento tecnologico" fosse in linea con le "capacità" della macchina.

In realtà non si tratta di una novità visto che già nel 2020 il legislatore, con la modifica dell'art. 3-*bis* nella L. 241/1990,[1] aveva posto la regola generale dell'utilizzo di strumenti informatici e telematici da parte delle PP.AA. *per conseguire maggiore efficienza nella loro attività, nei rapporti interni, tra le diverse amministrazioni e tra queste e i privati.*

L'introduzione dell'articolo 30 nel d.lgs. 36/2023 ha però aperto al tema dell'utilizzo degli algoritmi e dei sistemi di intelligenza artificiale nel più ampio spettro dell'attività amministrativa, in tal senso il primo comma prevede: *"per migliorare l'efficienza le stazioni appaltanti e gli enti […] provvedono, ove possibile, ad automatizzare le proprie attività ricorrendo a soluzioni tecnologiche, ivi incluse l'intelligenza artificiale […]."*

Quindi, la disciplina dell'art. 30 del Codice Appalti introduce l'uso di procedure automatizzate nel ciclo di vita dei contratti pubblici dimenticando che tutta la contrattualistica pubblica è permeata di verifiche e provvedimenti amministrativi vale a dire di procedure che avvengono unicamente su piattaforme digitali e sulle quali il singolo dipendente pubblico, (in genere l'ufficio di supporto al RUP (Responsabile Unico del Progetto) è chiamato ad operare attraverso un'auto formazione che deve essere immediata e continuativa.

Pertanto, mentre il grado di attendibilità e di ausilio in ambito professionale si va facendo strada nel mondo giurisprudenziale, il

[1] Articolo modificato con D.L. 76 del 16.07.2020, n. 76 che ha disposto (con l'art. 12, comma 1, lettera b)) la modifica dell'art. 3-bis, comma 1 della Legge 241/1990, "Uso della telematica": *per conseguire maggiore efficienza nella loro attività, le amministrazioni pubbliche ((agiscono mediante strumenti informatici e telematici)), nei rapporti interni, tra le diverse amministrazioni e tra queste e i privati.*

mondo degli appalti deve correre perché i fondi non aspettano la verifica sul grado di attendibilità e, soprattutto, la politica proclama "lavorazioni" e "cantierizzazioni" oggi per l'altro ieri.

Resta sempre e comunque, soprattutto nell'ambito della Pubblica Amministrazione, la necessarietà della competenza umana per verificare la correttezza, la veridicità e i combinati disposti del ragionamento composto ma non composito dalla macchina.

In sintesi, se l'articolo 30 del d.lgs. 36/2023 non pare così innovativo quel che sembra aver portata rivoluzionaria è l'applicazione concreta dell'intelligenza artificiale al panorama delle gare pubbliche.

Le sfide dell'intelligenza artificiale nella pubblica amministrazione italiana ruotano intorno a quattro pilastri: competenze, governance, sovranità tecnologica e qualità del dato. A dirlo sono gli stessi CTO e Data Manager dei nostri enti, come emerso anche dal Tavolo di CTO & Data Manager svoltosi nell'ambito di FORUM PA 2025, che ha promosso un confronto sull'evoluzione delle infrastrutture tecnologiche e sulla data governance in relazione all'adozione dell'IA nella pubblica amministrazione.[2]

È chiaro che questa innovazione ha ingenerato un vivace dibattito sia tecnologico che giuridico dando luogo a una serie di tematiche che riguardano due aspetti cruciali sui quali ci si soffermerà in questo capitolo:

il rapporto tra il contributo umano e la responsabilità dell'"operatore" privato e pubblico – ex lett. b), co. 3 dell'art. 30 d.lgs. 36/2023

la valutazione d'impatto come strumento principale di governance e i suoi relativi limiti.

Il brocardo *ubi societas, ibi ius*, dunque, nel rievocare il pensie-

[2] ForumPA 2025 dal 19 al 21 maggio, *Verso una PA aumentata – Persone - Tecnologie – Relazioni*, Roma, Palazzo dei Congressi.

ro di Santi Romano, si colora di nuove ed inattese suggestioni riconoscendo appieno un ruolo anche alla comunità che vive e alimenta una realtà caratterizzata dall'immaterialità.

2. Contributo umano e responsabilità dell'"operatore" privato e pubblico

Con l'articolo 30 il d.lgs. sancisce il principio della non esclusività della decisione algoritmica prevedendo che nel processo decisionale vi sia sempre un contributo umano capitalizzato attraverso una specifica attività di controllo, validazione, ovverosia della possibilità di smentita della decisione assunta con modalità automatizzate.[3] Tale principio, sintetizzato nell'elegante espressione "riserva di umanità", ha una duplice valenza: da una parte impone il contributo umano nell'adozione di una scelta, dall'altra espone il medesimo agente ad una forma di responsabilità rafforzata, onerandolo della necessità di acquisire a sua volta una piena conoscenza del funzionamento degli algoritmi utilizzati.[4]

Ma il focus è proprio sulle declinazioni del contributo umano per giungere, eventualmente, anche alla smentita della decisione automatizzata, ai sensi della lett. b) del co. 3 dell'art. 30 del Codice dei contratti pubblici che, come detto, impone la non esclusività della decisione algoritmica.

In effetti, tutto il corpus dell'AI Act dell'Ue è impregnato dalla necessità dell'intervento umano o, meglio, della "sorveglianza

[3] E. INGENITO, *L'automazione dei processi nei contratti pubblici*, in «Rivista Italiana di Informatica e Diritto», 2, 2025.

[4] G. GALLONE, *Riserva di umanità e funzioni amministrative. Indagine sui limiti dell'automazione decisionale tra procedimento e processo*, Wolters Kluwer, Milano, 2023.

umana",[5] finalizzata sia ad evitare e correggere le possibili distorsioni sistemiche sia all'adozione di «misure tecniche e organizzative per garantire che i sistemi di IA ad alto rischio siano quanto più possibile resilienti per quanto riguarda errori, guasti o incongruenze che possono verificarsi all'interno del sistema o nell'ambiente in cui esso opera, in particolare a causa della loro interazione con persone fisiche o altri sistemi» (art. 15.3 AI Act).

Ma l'AI Act dell'Ue contiene anche uno "spauracchio" non di poco conto che allarmerà le già sobillate stazioni appaltanti (in pratica, tutta la nostra Pubblica Amministrazione), vale a dire il considerando n. 96 che responsabilizza, in maniera massiva, gli "operatori",[6] i quali vengono individuati in "*qualsiasi persona fisica o giuridica, autorità pubblica, agenzia o altro organismo che utilizza un sistema di IA sotto la sua autorità, tranne nel caso in cui il sistema di*

[5] Cfr. considerando nn. 1, 4-*bis*, 9-*bis*, 43 e l'art. 4-*bis*. Tra gli altri si segnalano poi gli artt. 7 e 14 anche se riferiti alla possibilità «*di intervenire sul funzionamento del sistema di IA ad alto rischio o di interrompere il sistema mediante un pulsante di "arresto" o una procedura analoga, che consenta di arrestare il sistema in condizioni di sicurezza*».

[6] Considerando 96: «*Gli operatori di sistemi di IA ad alto rischio svolgono pertanto un ruolo fondamentale nel garantire la tutela dei diritti fondamentali, integrando gli obblighi del fornitore nello sviluppo del sistema di IA. Gli operatori sono nella posizione migliore per comprendere come il sistema di IA ad alto rischio sarà utilizzato concretamente e possono pertanto individuare potenziali rischi significativi non previsti nella fase di sviluppo* – omissis – *Gli operatori dovrebbero individuare strutture di governance adeguate in tale contesto specifico di utilizzo, quali le modalità di sorveglianza umana, le procedure di gestione dei reclami e le procedure di ricorso* – omissis – *Nell'effettuare tale valutazione d'impatto, l'operatore dovrebbe informare l'autorità nazionale di controllo e, nella misura più ampia possibile, i pertinenti portatori di interessi nonché i rappresentanti dei gruppi di persone che potrebbero essere interessati dal sistema di IA al fine di raccogliere le informazioni pertinenti ritenute necessarie per effettuare la valutazione d'impatto e sono incoraggiati a rendere pubblica la sintesi della loro valutazione d'impatto sui diritti fondamentali sul loro sito web online*».

IA sia utilizzato nel corso di un'attività personale non professionale".

A livello normativo, quindi, il pericolo più grave dell'intelligenza artificiale applicata ai contratti pubblici non riguarda il suo "acquisto" ma il suo concreto utilizzo poiché la relativa responsabilità è già sin da subito scaricata sull'operatore che, oltre a tutti gli obblighi previsti dalla copiosa normativa nazionale e ai doveri contenuti negli artt. 23, 27, 28, 29, 51 e 62 dell'AI ACT, ha anche il dovere di sovrintendenza umana sulle attività automatizzate come elencate dall' art. 27) dell'AI ACT.[7]

Il rischio viene riconnesso all'immanenza del suo uso ed è scaricato sull'operatore che viene individuato come responsabile e come destinatario di una notevole policromia di obblighi, tra i quali brilla il dovere di sovrintendenza umana sulle attività automatizzate. [8]

[7] Art. 27, A.I. Act: "*Prima di mettere in servizio un sistema di IA ad alto rischio quale definito all'articolo 6, paragrafo 2, ad eccezione dei sistemi di IA destinati ad essere utilizzati nel settore 2 dell'allegato III, gli operatori effettuano una valutazione dell'impatto dei sistemi nel contesto specifico di impiego …*".

[8] A.I. Act, combinato disposto artt. 13, 14, 26, 27 e Allegato IV- "Documentazione tecnica ex art. 11, paragrafo 1":

"*Nella misura in cui esercitano un controllo sul sistema di IA ad alto rischio, gli operatori:*

i) attuano la sorveglianza umana conformemente ai requisiti stabiliti nel presente regolamento;

ii) garantiscono che le persone fisiche preposte ad assicurare la sorveglianza umana dei sistemi di IA ad alto rischio siano competenti, adeguatamente qualificate e formate e dispongano delle risorse necessarie per assicurare l'efficace supervisione del sistema di IA a norma dell'art. 14;

iii) garantiscono che le misure pertinenti e adeguate in materia di robustezza e cibersicurezza siano periodicamente monitorate per verificarne l'efficacia e siano periodicamente adeguate o aggiornate".

Anche la Relazione illustrativa al d.lgs. 36/2023 conteneva il recepimento non soltanto dei principi sul tema dell'utilizzo di soluzioni di intelligenza artificiale ma anche dei principi enunciati dai giudici amministrativi con riguardo alle procedure di gara:

"Sono recepiti non soltanto i principi, che si sono da tempo affermati in ambito europeo sul tema dell'utilizzo di soluzioni di intelligenza artificiale ma anche […] quelli enunciati dai giudici amministrativi, secondo cui le stazioni appaltanti, in sede di acquisto o sviluppo delle soluzioni tecnologiche, si assicurano la disponibilità del codice sorgente e di ogni altro elemento utile a comprenderne le logiche di funzionamento; che la decisione assunta all'esito di un processo automatizzato deve considerarsi imputabile alla stazione appaltante. Sono altresì recepiti i principi per cui il processo decisionale non deve essere lasciato interamente alla macchina, dovendo assicurarsi il contributo umano per controllare, validare ovvero smentire la decisione automatica e per cui la decisione algoritmica non comporti discriminazioni di sorta. Il comma 4 prevede che le PP.AA. adottino ogni misura tecnica e organizzativa idonea a garantire che siano rettificati i fattori che comportano inesattezze dei dati e sia minimizzato il rischio di errori, nonché a impedire effetti discriminatori nei confronti di persone fisiche […]. Il comma 5, infine, prevede un adempimento per dare trasparenza e consentire ai cittadini e agli operatori di sapere se nelle attività riferite agli appalti pubblici si utilizzano o meno soluzioni tecnologiche avanzate […] imponendo alle pubbliche amministrazioni di darne contezza sul sito istituzionale, nella sezione "Amministrazione trasparente".[9]

Ne deriva che lo strumento dell'I.A. può essere sicuramente di grande aiuto per le operazioni ripetitive non caratterizzate da discrezionalità stanti i benefici attesi in termini di riduzione della tempistica procedimentale della P.A. ma le tecnologie più "appeti-

[9] Camera dei Deputati, *Relazione illustrativa al d.lgs. 36/2023*, 48, (art. 30).

bili", in termini di aumento dell'efficienza anche nell'ambito dell'affidamento degli appalti, restano probabilmente gli algoritmi conversazionali, quelli predittivi e quelli generativi di nuovi dati.

3. L'attendibilità del modello

Alla base dell'attendibilità di un modello di I.A., anche per la Pubblica Amministrazione, restano fondamentali la selezione corretta delle nozioni di base, la semplificazione e la chiarezza degli input che devono essere considerati dalla formula multidisciplinare che formerà l'algoritmo. Occorre poi integrare con il sistema di addestramento dell'Intelligenza Artificiale per condizionare il livello qualitativo dell'esito atteso e i data training attraverso cui il sistema viene addestrato. Si tratta di un set di dati qualitativamente e quantitativamente adeguato. Il legislatore comunitario nell'AI Act dell'UE affronta questa problematica precisando che un accesso ai dati di alta qualità svolge un ruolo essenziale per garantire prestazioni e pratiche di governance del sistema di IA adeguate, da sottoporre a convalida per evitare distorsioni che potrebbero provocare rischi per i diritti fondamentali o discriminazioni vietate.

Il primo presupposto per un'adeguata e legittima utilizzazione dell'intelligenza artificiale nei procedimenti di gara dovrebbe riguardare proprio l'appalto del servizio (o fornitura) del sistema di IA; vale a dire la necessità di reperire sul mercato un sistema supervisionato che sia stato addestrato in un ambiente adeguato affinché possa restituire output che siano il più possibile in linea o comunque che possano essere controllati, verificati e, nell'eventualità, anche smentiti da un "sorvegliante umano" (art. 14.1 AI Act).

In sintesi: occorre comprare il servizio (e forse anche la forni-

tura) che consentirà di istruire, determinare, pubblicare, aggiudicare, affidare, eseguire e liquidare un altro servizio (o lavoro/opera pubblica o fornitura).

In realtà questo già accade in tutte le PP.AA. italiane per la maggior parte dei servizi di progettazione, ingegneria ed esecuzione dei lavori pubblici: si compra qualcuno che progetti, ingegnerizzi ed esegua un lavoro/servizio/fornitura ancora da appaltare. In sintesi: si appalta l'omino (il professionista) che dovrà disegnare il quadro entro il quale dovrà operare (sulle nostre strade, per esempio) una ditta anche essa da appaltare.

4. L'imputazione della responsabilità

Quale deve essere il "giusto" contributo umano per giungere, eventualmente, anche alla smentita della decisione automatizzata, ai sensi della lett. b) del co. 3 dell'art. 30 del Codice dei contratti pubblici?

Entrando nel merito per individuare il centro di imputazione della responsabilità per un danno causato da un errore dell'IA la fattispecie diventa complessa poiché l'IA è una "*mera res*" e, come tale, non sanzionabile. Si dibatte su chi debba rispondere: il produttore, il programmatore, l'utilizzatore o la pubblica amministrazione. Il Regolamento europeo sull'IA (AI Act) ha tentato di armonizzare il regime di responsabilità ma permangono i dubbi.

Nonostante i notevoli vantaggi in termini di efficienza, trasparenza e riduzione dei *bias* cognitivi che l'IA può apportare agli appalti pubblici, permangono sfide significative. L'opacità dei sistemi di *machine learning*, il rischio di "allucinazioni" e la complessa questione della responsabilità per i danni causati dagli algoritmi

richiedono un continuo monitoraggio e adattamento del quadro normativo e giurisprudenziale.[10]

Si richiede, pertanto, che i sistemi utilizzati siano sottoposti alla sorveglianza umana imponendo che i soggetti preposti siano adeguatamente formati e con competenze tali da garantire un'efficace supervisione dei sistemi di intelligenza artificiale utilizzati, un'attività di prevenzione con riguardo alla cybersicurezza, nonché di monitoraggio per valutare il sistema in termini di adeguatezza ed efficacia, oltreché per individuare i correttivi richiesti per procedere ad un aggiornamento periodico.

Tale gravoso onere, che per taluni osservatori – in virtù dei profili di responsabilità che ne derivano – potrebbe semplicemente determinare la fuga delle stazioni appaltanti dal ricorso alle procedure automatizzate, tiene conto della circostanza che i pericoli maggiori scaturenti dal ricorso all'intelligenza artificiale non derivano dalla sua produzione e messa in commercio, ma dal suo utilizzo.[11]

5. Rischi e Responsabilità

Un adeguato sistema di governance rappresenta il secondo pilastro per la migliore e legittima utilizzazione nel campo dei pubblici appalti dei sistemi di intelligenza artificiale.

Alcuni studiosi statunitensi hanno evidenziato la specificità e l'assoluta novità della sfida lanciata dall'intelligenza artificiale che

[10] A. MORETTI, *Intelligenza artificiale e contratti pubblici: la distinzione tra algoritmi di supporto e decisionali - L'accesso al codice sorgente e il bilanciamento fra trasparenza e proprietà intellettuale*, in «Altalex», 17.07.2025.

[11] E. INGENITO*, L'automazione dei processi nei contratti pubblici*, in «Rivista Italiana di Informatica e Diritto», 2, 2025.

impone risposte totalmente nuove e altre rispetto ad una concezione di controllo e verifica ancora tradizionale.

La filosofia del Regolamento (UE) 2024/1689 è quella di un approccio preventivo basato sul rischio che stabilisce obblighi per fornitori e operatori dei sistemi di IA a seconda del livello di rischio che questi possono comportare per la salute, la sicurezza e i diritti fondamentali delle persone. Tra questi rischi l'*AI Act* classifica i sistemi di IA in quattro categorie di rischio, tra le quali compare proprio la questione de qua:

i sistemi di IA ad alto rischio sono consentiti ma con numerosi obblighi (gestione dei rischi, governance e gestione dei dati, valutazioni di conformità, trasparenza, misure di sicurezza, supervisione umana, documentazione dettagliata). Sebbene il settore degli appalti pubblici non sia qualificato *ex se* come ad alto rischio, alcune specifiche applicazioni di IA al suo interno potrebbero rientrare in tale categoria;[12]

Al fine di prevenire eventuali riscontri asseritamente discriminatori o comunque errati, limitati e parziali non bisogna soffermarsi sulla tipologia di algoritmo adoperato, bensì sull'ambiente e sui dati che al momento dell'addestramento possano ritenersi sufficientemente adeguati in termini quantitativi e qualitativi.

Il secondo presupposto indefettibile evoca, invece, la predisposizione di un sistema di governance adeguato o, meglio, di presidi volti ad effettuare dei controlli e delle verifiche penetranti sia *ex ante* che *ex post*. Un presidio preventivo efficace consente, del resto, di favorire le condizioni ottimali per addestrare il sistema in un ambiente adeguato.[13]

[12] A. MORETTI, *op. cit.*

[13] F. CARINGELLA, L. CARBONE, G. ROVELLI, *Manuale dei contratti pubblici*, Dike Giuridica, Napoli, 2024.

Nello specifico si ritiene che al soggetto responsabile della gestione del meccanismo di automazione spetti altresì pianificare l'organizzazione di strutture *ad hoc* con cui attuare gli obblighi di sorveglianza umana, nonché gestire, in primis, la formazione e poi "inventare" un sistema per tutto il contenzioso.

La prospettiva, in particolare, è duplice:

in un'ottica preventiva è necessario predisporre un piano di valutazione dettagliato che, nel raccogliere le informazioni necessarie e nel descrivere l'impatto che deve avere il sistema di IA sotto il profilo qualitativo e quantitativo, individui altresì le misure e gli strumenti per mitigare i rischi con particolare riferimento ai diritti fondamentali astrattamente pregiudicati dal concreto uso del predetto sistema;

in un'ottica di controllo/verifica occorre elaborare un presidio per il monitoraggio continuo volto a misurare l'adeguatezza e l'efficacia del sistema nel tendenziale rispetto dei principi esposti in precedenza.

La necessità di valutare, in modalità preliminare, l'impatto sul sistema concretamente utilizzato in relazione al contesto in cui opera è un modello che sul piano europeo ha già trovato applicazione in altri settori sensibili. Si pensi alla direttiva 85/337/CEE in ordine alla valutazione d'impatto ambientale, all'art. 35 del Regolamento (UE) 2016/679 sulla valutazione d'impatto sulla protezione dei dati derivante dal ricorso alle nuove tecnologie, ma anche alle tecniche proprie della prevenzione e del contrasto alla corruzione.[14]

In sostanza, il sistema di governance dovrebbe consentire di individuare preventivamente i rischi e predisporre i meccanismi

[14] U. RUFFOLO, *XXVI lezioni di Diritto dell'intelligenza artificiale,* Giappichelli, Torino, 2021.

per limitare gli effetti pregiudizievoli e le conseguenze negative derivanti dal sistema di intelligenza artificiale.

Sebbene sia del tutto illusorio auspicare di poter irreggimentare, entro schemi predefiniti, ogni rischio scaturente dallo sviluppo delle nuove tecnologie (in particolare con riguardo ai sistemi generativi di intelligenza artificiale), è indubbio che – pur nella consapevolezza che si tratti di una fenomenologia che, al pari del moto ondoso, non sia governabile – esse rappresentino un'opportunità irrinunciabile.

È necessario, per tale ragione, chiarire quali obiettivi possa perseguire la pubblica amministrazione in termini di imparzialità e buon andamento ed investire, come si è diffusamente sottolineato, in un sistema di governance che individui adeguati presidi in grado di intervenire tempestivamente sia in via preventiva che successiva su eventuali errori o anomalie. L'osservanza dei principi su cui si fonda l'azione amministrativa, del resto, non può che collimare con la "riserva di umanità" che deve ispirare tutto lo sviluppo delle nuove tecnologie.

Ciò, tuttavia, impone un duplice sforzo. In primo luogo, è necessario investire nella formazione del personale per fornire competenze adeguate in grado di padroneggiare i sistemi più evoluti e di coglierne, almeno potenzialmente, i rischi che derivano da un dataset incompleto o basato su elementi erronei in grado di alimentare ipotesi di discriminazione algoritmica.

In secondo luogo, è imprescindibile riflettere sul modello su cui strutturare un sistema di certificazioni che possa consentire all'agente umano di effettuare in maniera incisiva una penetrante attività di verifica e di controllo.[15]

[15] E. INGENITO, *op. cit*.

6. L'illusione dell'imparzialità computazionale nella pubblica amministrazione

Nel dibattito contemporaneo sul rapporto tra intelligenza artificiale e pubblica amministrazione si è affermata una descrizione tanto seducente quanto insidiosa: "l'algoritmo come garante ultimo dell'imparzialità amministrativa" e, di conseguenza, il riconoscimento "orwelliano" della natura amministrativa del "provvedimento algoritmico"[16]. La promessa è quella di una giustizia procedurale finalmente emancipata dalle debolezze umane (corruzione, negligenza, inerzia, arbitrio etc.). Un sistema computazionale,[17] si argomenta, non conosce favoritismi, non subisce pressioni, non cede a tentazioni. In questa prospettiva, l'automazione decisionale si configurerebbe come il compimento tecnico (paradossale?) del principio costituzionale di imparzialità.

Tuttavia, questa visione riposa su un equivoco fondamentale che merita un'analisi gius-filosofica approfondita: non solo la confusione tra imparzialità formale e neutralità sostanziale ma tra efficienza procedurale e giustizia materiale.

L'impiego di algoritmi sarà certamente in grado di valutare e graduare una moltitudine di domande soddisfacendo le esigenze di celerità dell'azione amministrativa ma, ad oggi, non può essere

[16] https://www.irpa.eu/provvedimento-algoritmico-si-ma-come/ in «Osservatorio sullo Stato digitale».

[17] È il tema della qualità del risultato, che, nelle decisioni automatizzate, s'interseca strettamente a quello dei *Big Data*, i quali, per definizione, si fondano su un concetto prettamente quantitativo piuttosto che qualitativo. Sui *Big Data* si veda P. DELORT, *Le big data*, PUF, Parigi, 2015; A. OTTOLIA, Big data *e innovazione computazionale*, Giappichelli, Torino, 2017; C. COMELLA, *Origine dei* "Big Data", in «Gnosis», 2, 2017, 130; in senso critico, v. G. AVANZINI, *Intelligenza artificiale*, machine learning *e istruttoria procedimentale: vantaggi, limiti ed esigenze di una specifica* data governance, in A. PAJNO, F. DONATI, A. PERRUCCI, *Intelligenza artificiale e diritto: una rivoluzione?,* Il Mulino, Bologna, 2022, 75 ss.

scollegato dalle tutele che il nostro ordinamento deve riconoscere anche al soggetto destinatario degli effetti giuridici di una eventuale determinazione non umana.[18]

Come osservava già il Consiglio di Stato in una pronuncia del 2019 l'utilizzo di tecniche algoritmiche può garantire «*l'esclusione di interferenze dovute a negligenza (o peggio dolo) del funzionario»,* realizzando così una «*maggior garanzia di imparzialità della decisione automatizzata*». Ma questa imparzialità – è lecito chiedersi – corrisponde davvero all'ideale costituzionale che informa l'azione amministrativa?

6.1. L'illusione epistemologica: il dato come costruzione sociale

Per comprendere i limiti strutturali dell'imparzialità algoritmica occorre innanzitutto decostruire un mito epistemologico: quello della neutralità del dato. Come ci ricorda la riflessione critica sulla datificazione, il dato non è mai semplicemente "dato" (*datum*), ma è sempre "da dare" (*datur)* – frutto di scelte metodologiche, categorie interpretative, decisioni circa ciò che merita di essere osservato e ciò che può essere trascurato.

I dati sono il risultato di un'indagine e non la sua premessa. Non esistono in natura come entità autodeterminate, "fresche" cioè pronte per essere "raccolte", ma vengono prodotti attraverso processi che incorporano necessariamente prospettive culturali, priorità politiche e pregiudizi cognitivi. Ogni dataset riflette e riporta le condizioni storiche e sociali della sua generazione, cristallizzando in forma numerica le relazioni di potere, le esclusioni sistemiche, le gerarchie e i vizi impliciti della società che lo ha prodotto.

Quando, dunque, un algoritmo viene addestrato su dati storici

[18] https://www.irpa.eu/provvedimento-algoritmico-si-ma-come/.

relativi, ad esempio, all'aggiudicazione di appalti pubblici in contesti caratterizzati da pratiche corruttive endemiche, non può miracolosamente "epurare" quei dati dalle loro distorsioni. Al contrario, il sistema apprenderà proprio quelle patologie come regolari trasformandole in parametri predittivi e, in ultima analisi, normalizzandole come standard di riferimento. La presunta neutralità computazionale si rivela così una forma particolarmente insidiosa di conservatorismo epistemologico: l'algoritmo non elimina il *bias*, lo rende invisibile, lo traveste da oggettività matematica e, quindi, provvedimentale.

6.2. L'etica computazionale: la responsabilità tra scelta morale e calcolo automatico

Ma la questione più radicale che l'automazione decisionale pone al diritto pubblico è di natura etica prima ancora che tecnica: può un sistema privo di coscienza morale assumersi la responsabilità di scelte che incidono sui diritti e sugli interessi dei cittadini?

Come ci ammonisce la riflessione sull'etica computazionale,[19] un algoritmo può ottimizzare una funzione obiettivo ma non può interrogarsi sul senso di quella funzione. Non conosce l'ambiguità, la contraddizione, il conflitto interiore che caratterizza ogni autentica scelta morale. L'intelligenza artificiale è, in senso proprio, premorale: non sceglie, calcola. Non delibera, applica.

Questa distinzione non è un sofisma filosofico ma ha conseguenze giuridiche decisive e, soprattutto, responsabilità cogenti.

[19] Si vedano in proposito: M. DURANTE, *Il futuro del web: etica, diritto, decentramento*, Giappichelli, Torino, 2007; A. FABRIS, *Etica per le tecnologie dell'informazione e della comunicazione*, Carocci Editore, Roma, 2018; F. CASA, S. GAETANO E G. PASCALI (a cura di), *Intelligenza artificiale: diritto, etica e democrazia*, Il Mulino, Bologna, 2025.

La responsabilità amministrativa – nelle sue molteplici declinazioni: civile, penale, disciplinare, contabile, contrattuale etc. – presuppone sempre un soggetto capace di intendere e di volere, un agente che possa essere chiamato a rispondere delle proprie scelte. Come può configurarsi un regime di responsabilità quando il decisore è un sistema algoritmico, una macchina?

La soluzione formale – mantenere un funzionario umano che "supervisiona" o "convalida" le decisioni automatizzate – si rivela spesso illusoria. L'automazione applicata alla cosa pubblica, sin dalle analisi di Jeremy Rifkin,[20] ha prodotto una sorta di apatia decisionale: l'essere umano, progressivamente disabituato all'esercizio del giudizio critico, finisce per limitarsi a ratificare ciò che la macchina ha già determinato. Delegando sempre più, pensa sempre meno. La responsabilità formale resta, ma la responsabilità sostanziale evapora.

6.3. Efficienza versus *giustizia: antica dicotomia tra principi e regole applicabili*

Un secondo argomento frequentemente invocato a sostegno dell'automazione amministrativa è quello dell'efficienza. L'intelligenza artificiale promette velocità, precisione, economicità: valori non certo estranei al principio costituzionale di buon andamento.

[20] J. RIFKIN, *The End of Work,* G. P. Putnam's Sons, New York, I ed. 1995, (tr. it. di Paolo Canton, *La fine del lavoro*, Milano, Baldini & Castoldi, IV ed. 1995, 18): *Le tecnologie dell'informazione e della comunicazione e le forze del mercato globale stanno rapidamente polarizzando la popolazione mondiale in due forze inconciliabili e potenzialmente conflittuali: una élite cosmopolita di 'analisti di simboli' che controllano le tecnologie e le forze di produzione; e un crescente numero di lavoratori permanentemente in eccesso, con poche speranze e ancor meno prospettive di trovare un'occupazione significativa nella nuova economia globale ad alta tecnologia.*

Ma è qui che riaffiora una dicotomia classica del pensiero giuridico: il rapporto, spesso tensionale, tra efficienza e giustizia.

L'utilizzo di tecnologie avanzate in ambito amministrativo risulta conforme ai canoni di efficienza e di economicità dell'azione amministrativa (art.1, l. 241/1990), i quali, secondo il principio costituzionale di buon andamento dell'azione amministrativa, impongono all'amministrazione il conseguimento dei propri fini con il minor dispendio di mezzi e risorse e attraverso lo snellimento e l'accelerazione dell'iter procedimentale. Il principio di legalità alla base del diritto amministrativo, visti i tempi, va riguardato come principio di legalità algoritmica che trova il suo principale fondamento proprio nell'art. 97 della Costituzione, in quanto l'introduzione dei modelli decisionali automatizzati e l'utilizzo di tecnologie informatiche risponde alla doverosa declinazione della norma costituzionale, coerente con l'attuale e futura evoluzione tecnologica.

Con riferimento all'attività amministrativa è noto che già a partire dal 2017[21] si è posto il tema della c.d. decisione algoritmica come quella assunta all'esito di procedure amministrative automatizzate che fanno applicazione di algoritmi. In particolare, è stata puntualizzata la necessità che essa soggiaccia ai principi previsti dalla legge sul procedimento amministrativo e, segnatamente, ai principi di imparzialità, pubblicità e trasparenza. Infatti, per quanto indubbi siano i vantaggi che derivano dall'automazione del processo decisionale dell'amministrazione mediante l'utilizzo di una procedura digitale e algoritmica, l'utilizzo della stessa non dovrebbe mai implicare un sacrificio dei principi fondamentali che segnano il procedimento amministrativo e che si sono oramai consolidati

[21] Ci si riferisce alla sentenza n. 3769 del 22.03.2017 con la quale il TAR Lazio, Sez. III-bis, aprì nuovi scenari in merito al diritto di accesso all'algoritmo.

nel nostro ordinamento (Cons. Stato, VI, 8 aprile 2019, n. 2270, 13 dicembre 2019, n. 8474).

Il ricorso all'algoritmo appartiene alla scelta organizzativo/discrezionale dell'amministrazione al fine di realizzare finalità di interesse pubblico e può ben essere utilizzato non solo per lo svolgimento dell'attività amministrativa vincolata ma anche per quella discrezionale, considerate entrambe espressione del potere autoritativo esercitato per il perseguimento del pubblico interesse. Ma la "scelta organizzativa" è fatta dai dipendenti, dai funzionari, dai dirigenti vale a dire dalle "intelligenze umane" della P.A. che deve affrontare una rivoluzione prima di tutto culturale. Nonostante le amministrazioni pubbliche siano pronte a sperimentare con l'intelligenza artificiale manca loro una chiara visione degli esiti finali di questi progetti. La vera sfida non è tecnologica, ma culturale: servono capacità specialistiche per gestire strumenti, processi e data governance. L'introduzione dell'IA richiede nuove figure professionali, ancora rare nella PA, rendendo indispensabile un piano strategico per attrarre e formare talenti capaci di gestire dati complessi e garantirne la qualità.

7. Attività amministrativa e algoritmi

Pur nella consapevolezza dell'indubbio "fascino" che accompagna lo studio dell'impiego nell'attività amministrativa di algoritmi, di apprendimento e non, in particolare per le evidenti ricadute in termini di efficienza e di velocizzazione della decisione pubblica, nonché del fatto che tale utilizzo rappresenta inevitabilmente il futuro delle procedure amministrative, nondimeno le amministrazioni dovranno prestare grande attenzione nell'assicurare il pieno rispetto dei diritti e delle prerogative dei soggetti

coinvolti, affinché l'era digitale non implichi un arretramento sul fronte dei diritti partecipativi e conoscitivi. Sarà, quindi, sempre necessario, in particolare, assicurare un'adeguata motivazione e trasparenza finalizzate a consentire la comprensibilità del "risultato algoritmico".

Sul fronte pubbliche amministrazioni e tutela degli interessi dei cittadini, tre sono i principi importanti e altamente sfidanti da considerare per avere un'IA effettivamente al servizio del cittadino, per consentirne un uso etico e portare fiducia in questa tecnologia, principi cristallizzati nel Regolamento UE 2024/1689 del 13 giugno 2024 (AI Act): assicurare che venga garantita la trasparenza dei sistemi di intelligenza artificiale in modo da comprenderne il funzionamento (art. 13), assicurare il diritto alla spiegabilità dei singoli processi decisionali in favore degli interessati (art. 86), e assicurare la sorveglianza umana (art. 14).

Come la storia ci insegna, anche i regimi più oppressivi possono essere "efficienti" nel perseguire i propri obiettivi. L'efficienza è neutrale rispetto ai fini: può servire il bene comune come può consolidare l'ingiustizia. Un algoritmo può assegnare risorse o aggiudicare contratti con straordinaria rapidità, ma non può valutare se quelle decisioni realizzano equità distributiva, se rispettano la dignità delle persone coinvolte, se servono autenticamente l'interesse pubblico.

Il diritto amministrativo moderno non è solo diritto dell'efficienza, ma anche – e soprattutto – diritto dei valori: imparzialità, trasparenza, partecipazione, proporzionalità. Questi principi non possono essere ridotti a meri parametri di ottimizzazione. Richiedono ponderazione, bilanciamento e contestualizzazione: operazioni intrinsecamente umane che nessun algoritmo può surrogare e che, molto spesso, risultano dubbie anche per l'intelligenza umana sia pubblica che privata.

8. Il nodo della trasparenza: l'opacità del "decision-making" automatizzato

Un ulteriore paradosso "dell'amministrazione algoritmica" riguarda la trasparenza. I sistemi di intelligenza artificiale, specialmente quelli basati su reti neurali profonde, operano secondo logiche che sfuggono spesso alla comprensione dei loro stessi creatori. Il fenomeno della *black box decisionale*[22] solleva interrogativi cruciali dal punto di vista del diritto amministrativo.

Come può un cittadino esercitare il proprio diritto di difesa contro una decisione automatizzata di cui non comprende le ragioni? Come può l'amministrazione motivare adeguatamente un provvedimento quando il percorso logico che ha condotto a quella determinazione è computazionalmente opaco? Come si concilia l'obbligo di trasparenza con l'inscrutabilità degli algoritmi?

Ad oggi, tra le possibili soluzioni adottabili, emergono le tecniche di interpretabilità post-hoc, che sono applicate successivamente alla produzione dell'output da parte del modello. Questo approccio analizza il comportamento dell'algoritmo, evidenziando quali variabili e processi abbiano maggiormente influenzato la decisione finale. Ad esempio, strumenti come LIME *(Local Interpretable Model-agnostic Explanations*) operano come "lenti d'ingrandimento" su singoli casi.

Il legislatore europeo ha prontamente riconosciuto l'impor-

[22] Per black-box decisionale si intendono quei sistemi la cui logica decisionale risulta opaca e difficilmente interpretabile dagli utenti. In proposito, cfr.: M. CARABANTES, *Black-box artificial intelligence: an epistemological and critical analysis*, in «AI & Society», 2, 2020; B. GOODMAN, S. FLAXMAN, *EU regulations on algorithmic decision-making and a "right to explanation",* ICML workshop on human interpretability in machine learning, New York, 2016; C. ZEDNIK, *Solving the black box problem: A normative framework for explainable artificial intelligence,* in «Philosophy & Technology», 2, 2021.

tanza della trasparenza nei sistemi di IA, in particolare già con l'introduzione del Regolamento Generale sulla Protezione dei Dati (GDPR), che prevede il diritto per gli utenti di non essere soggetti a decisioni basate unicamente su trattamenti automatizzati (art. 22). Tale principio impone che laddove vi siano giudizi o scelte demandate a processi algoritmici idonei ad incidere significativamente sulla vita degli individui, questi ultimi possano richiedere l'intervento umano ovvero ricevere appropriate spiegazioni. Al GDPR si affianca, ovviamente, il Regolamento sull'Intelligenza Artificiale, il quale classifica i sistemi di IA in base al rischio e prevede requisiti stringenti di trasparenza, soprattutto per le applicazioni che potrebbero influire sui diritti fondamentali o sulla sicurezza delle persone. Tali interventi normativi non solo forniscono linee guida essenziali per gli sviluppatori, ma rafforzano anche la fiducia del pubblico nei confronti delle nuove tecnologie, ponendo l'accento sulla necessità di un'IA responsabile e verificabile.

9. Appalti Digitali, al via la Rivoluzione 2.0: AgID vara le nuove regole tecniche per le piattaforme

Il Codice dell'amministrazione digitale riconosce alle pubbliche amministrazioni il potere-dovere di utilizzare le tecnologie dell'informazione per realizzare obiettivi di efficienza ed economicità, ma sempre nel rispetto dei principi di trasparenza e partecipazione. Ma quale trasparenza è possibile quando il processo decisionale si consuma in strati sovrapposti di funzioni matematiche non interpretabili?

Ed è infatti di questi giorni la Determinazione numero 267 del 31 dicembre 2025 con la quale l'AgID ha ufficialmente dato il via libera alla versione 2.0 delle Regole Tecniche per le Piattaforme di

Approvvigionamento Digitale (PAD) che dovranno essere implementate per rendere le gare d'appalto ancora più semplici, sicure e trasparenti.

Si tratta di un documento che contiene quattro allegati: (*i*) Regole tecniche - Requisiti tecnici e modalità di certificazione delle Piattaforme di Approvvigionamento Digitale (PAD); (*ii*) Modello di interoperabilità per le Piattaforme di approvvigionamento digitale, All. 1 al documento Regole tecniche; (*iii*) Checklist per la certificazione delle PAD, Allegato 2 al documento "Regole tecniche"; (*iv*) Quadro di riferimento normativo e tecnico per le Piattaforme di approvvigionamento digitale, Allegato al documento "Regole tecniche".

Le nuove regole, che sostituiscono le precedenti – adottate con la Determinazione AgID n. 137/2023 – evidenziano l'importanza della sicurezza informatica e della protezione dei dati, introducendo criteri più stringenti per rendere le infrastrutture digitali più robuste e resilienti.

Un'altra importante novità riguarda l'apertura verso le tecnologie emergenti: AgID lancia le prime linee guida per l'integrazione dell'Intelligenza Artificiale nei processi di gara, sempre sotto il controllo umano e nel rispetto della massima trasparenza.

Tra le Regole Tecniche vi è un intero capitolo sugli ambiti applicativi potenziali nelle PAD che accenna proprio alle funzionalità basate sull'IA in veste di supporto per le stazioni appaltanti e per gli operatori economici elencandone il relativo impiego con alcuni esempi:

supporto alla progettazione di gara: analisi predittiva dei rischi, suggerimenti per la definizione dei criteri di valutazione, verifica automatica della completezza e coerenza della documentazione di gara:

analisi e valutazione delle offerte: classificazione automatica dei documenti presentati dagli operatori economici, estrazione e confron-

to di dati strutturati, supporto alla valutazione degli aspetti qualitativi delle offerte tecniche (identificando elementi rilevanti o pattern specifici), individuazione di potenziali offerte anomale sulla base di indicatori predefiniti:

gestione della procedura: automazione di comunicazioni standardizzate, monitoraggio delle scadenze, supporto alla gestione del contenzioso tramite analisi della documentazione:

fase esecutiva: monitoraggio avanzato dell'esecuzione contrattuale, analisi predittiva di possibili ritardi o criticità, supporto alla gestione della contabilità e dei pagamenti:

elaborazioni statistiche e analisi predittive rispetto a procedure di approvvigionamento simili, allo scopo di ottimizzare processi e trovare forme di potenziali aggregazioni.

Negli annunci l'Agenzia governativa spiega che non si tratta di un semplice aggiornamento burocratico ma di un vero e proprio "upgrade di sistema" che punta a rendere le gare d'appalto più veloci, blindate sul fronte della sicurezza e, per la prima volta, aperte all'uso regolamentato dell'Intelligenza Artificiale.

In estrema sintesi è stato ridefinito il perimetro tecnologico entro cui dovranno muoversi stazioni appaltanti e operatori economici inaugurando una nuova stagione per l'e-procurement.

10. Esperienze applicative: tra sperimentazione e cautela

Alcune amministrazioni hanno intrapreso la via dell'automazione nel controllo degli appalti pubblici e nella prevenzione della corruzione. Queste sperimentazioni – condotte con gradi variabili di ambizione e consapevolezza critica – offrono un terreno prezioso per verificare empiricamente le promesse e i limiti dell'intelligenza artificiale applicata alla funzione di controllo.

Ed è infatti proprio il supremo giudice amministrativo che in una sentenza del 4 giugno 2025 ammoniva la PA dall'utilizzo esclusivo tramite algoritmi nelle procedure di gara.[23]

Tuttavia, è essenziale distinguere tra utilizzi diversi della tecnologia. Una cosa è impiegare sistemi di *data analytics* per identificare anomalie o *pattern* sospetti che meritano un approfondimento da parte di funzionari umani; altra cosa è delegare all'algoritmo la decisione finale, trasformandolo da strumento ausiliario in sostituto del decisore pubblico.

La prima soluzione – quella dell'algoritmo come supporto alla decisione umana – appare compatibile con i principi costituzionali, a condizione che vengano garantite: la verificabilità dei criteri utilizzati, la possibilità di contestare le segnalazioni automatiche, la formazione adeguata del personale chiamato a interpretare i risul-

[23] Consiglio di Stato, sent. n. 4857 del 4 giugno 2025: *L'accesso al codice sorgente della piattaforma di e-procurement è consentito solo se è l'unico mezzo di prova per la difesa in giudizio dei propri interessi*. Si tratta di una questione di cruciale importanza: l'utilizzo degli algoritmi nelle procedure di gara pubbliche. Per la IV Sezione del Consiglio di Stato, ai sensi dell'art. 35, comma 5, del Codice dei contratti, «*La nozione di "indispensabilità", cui il codice appalti subordina l'accesso nell'ipotesi sopra descritta, deve essere declinata nel senso di insussistenza di altri mezzi di prova idonei a dimostrare i fatti oggetto di contesa tra le parti. La semplice volontà di verificare e sondare (come è nel caso di specie) non legittima un accesso "meramente esplorativo" alle informazioni riservate, in quanto difetterebbe la comprova della specifica e concreta indispensabilità ai fini difensivi*»; ne consegue che in caso di utilizzo da parte della S.A. di una piattaforma digitale coperta da diritti di privativa intellettuale, l'istanza del concorrente escluso (per la mancata corrispondenza tra l'impronta digitale di un documento trasmesso con quella assegnata dal sistema in fase di conferma di partecipazione) di accesso difensivo ex art. 35, co. 5, d.lgs. 36/2023 al "codice sorgente" non può essere accolta se la richiesta è motivata semplicemente con la necessità di verificare il corretto funzionamento del sistema ed il richiedente non fornisca almeno un principio di prova in ordine al malfunzionamento della piattaforma. Ciò equivale ad affermare che la decisione algoritmica è legittima solo se gli algoritmi sono a supporto.

tati. La seconda soluzione – quella della piena automazione decisionale – solleva invece dubbi che non possono essere risolti mediante espedienti terminologici o aggiustamenti normativi marginali.

11. Sperimentazioni normative: l'hub dei contratti pubblici

Nell'aprile del 2025 il Ministero delle Infrastrutture e dei Trasporti (MIT) ha introdotto un nuovo strumento per semplificare le procedure operative delle Stazioni Appaltanti partorendo l'hub dei contratti pubblici[24], una piattaforma di riferimento per la contrattualistica pubblica con un'unica interfaccia nella quale si trovano il Codice aggiornato e commentato coi pareri del Servizio Giuridico, i corsi di formazione, i modelli di documenti, gli iter da seguire ('Bussola del RUP'), giurisprudenza e molto altro.

Questa hub è continuamente implementata e, infatti, l'ultima novità riguarda la sezione "digitApp" che, oltre a contenere il codice dei contratti pubblici digitalizzato dal 1° gennaio 2026 dispone di una nuova funzionalità di D*ocument Composition* con lo scopo di supportare il personale delle SA e centrali di committenza nella gestione efficace del processo di approvvigionamento pubblico. Il sistema di *Document Composition* automatizza la generazione documentale con elevata precisione, ottimizzando tempi e minimizzando errori.

La sezione contiene anche Chatbot DigitAPP in fase di sperimentazione su ambiti territoriali differenti, progettati per testare le funzionalità del sistema e raccogliere dati, utili a migliorarne prestazioni ed efficacia, un assistente digitale intelligente, progettato

[24] https://www.serviziocontrattipubblici.it.

per assicurare più efficienza, coerenza normativa e facile accesso alle informazioni del mondo dei contratti pubblici.

Altra novità di rilievo è l'elaborazione e la valorizzazione della giurisprudenza amministrativa in materia di contratti pubblici. Attraverso l'estrazione e la strutturazione di dati provenienti da sentenze e pareri pubblicati, in particolare, sul portale OpenGA,[25] verranno sviluppate funzionalità avanzate di consultazione, ricerca semantica e supporto alle decisioni. Saranno previsti, inoltre la creazione di un Codice dei contratti pubblici digitale annotato con pareri, giurisprudenza, annotazioni, costantemente aggiornato, nonché l'utilizzo dell'Al generativa e non generativa per sfruttare, adattare o arricchire i documenti introdotti nel sistema.

12. Conclusioni: per una filosofia della responsabilità nell'era algoritmica

L'intelligenza artificiale rappresenta indubbiamente un'opportunità straordinaria per il miglioramento dell'azione amministrativa. Sarebbe miope rifiutarla per preconcetto luddista. Ma sarebbe altrettanto irresponsabile abbracciarla acriticamente cedendo alla seduzione della neutralità computazionale.

In definitiva, portare l'IA nella PA italiana richiede un approccio integrato "multifattoriale" perché l'implementazione sostenibile e sicura ha bisogno della compresenza di competenze, governance, dati, infrastrutture, sicurezza ed ecosistema.

La trasformazione della PA tramite l'intelligenza artificiale esige investimenti infrastrutturali, un forte impegno nel cambiamento culturale e nell'acquisizione di competenze e, soprattutto, una vi-

[25] Open GA - Open Data è il nuovo portale degli open data di Giustizia Amministrativa attivato nel 2024: https://openga.giustizia-amministrativa.it.

sione strategica che promuova la collaborazione e la creazione di ecosistemi condivisi a livello nazionale ed europeo. Solo così la PA potrà sfruttare appieno il potenziale dell'IA a beneficio dei cittadini e del Paese.

Sfruttare la risorsa dell'I.A. nella PA è una sfida da non perdere per agevolare risposte tempestive ed essere più vicini ai cittadini. Si stima che essa genererà un aumento di produttività notevole soprattutto nelle politiche e nei servizi pubblici, nelle funzioni di supporto, nella semplificazione amministrativa e normativa, nelle attività delle Agenzie e dei dipartimenti statali. Ciò, tenuto conto del rapido invecchiamento della forza lavoro che mette a rischio la competitività del sistema Italia. Va poi rilevata l'esorbitante sovrapposizione di competenze richieste al dipendente pubblico che, se fino a ieri, per operare nel campo degli appalti, necessitava di possedere un variegato spettro di competenze a valenza giuridica, ad oggi sarebbero richieste anche elevate prestazioni di natura informatica e tecnologica.

Ciò di cui abbiamo bisogno non è un'etica "dell'IA" – espressione che rischia di essere un ossimoro – ma un'etica per l'IA: una riflessione rigorosa sui limiti entro cui la delega algoritmica può essere considerata legittima, sui meccanismi di controllo democratico che devono presidiarla, sulle forme di responsabilità che devono accompagnarla.

Come osservava Einstein, «*la morale non ha niente di divino, è una faccenda puramente umana*». E Nietzsche ci ricordava che ogni sistema morale è storicamente situato, culturalmente determinato. L'etica algoritmica, se non vogliamo che diventi una "morale della macchina" – vuota, funzionale, sterilizzata – deve restare saldamente ancorata alla responsabilità umana.

Il principio di imparzialità amministrativa, sancito dall'articolo 97 della Costituzione, non può essere ridotto a coerenza computa-

zionale ma richiede qualcosa di più profondo: la capacità di ascolto, il discernimento del contesto, l'empatia verso le situazioni concrete, la consapevolezza delle conseguenze delle proprie scelte. Qualità irriducibilmente umane.

L'algoritmo può vigilare, ma non può giudicare. Può segnalare, ma non può decidere. Può assistere, ma non può sostituire. Il confine tra questi utilizzi è il confine stesso tra amministrazione e automatismo, tra democrazia e tecnocrazia, tra diritto e calcolo.

Non temiamo abbastanza quando l'intelligenza artificiale sbaglia. Temiamo troppo poco quando non ci chiede più di pensare.

Sanità digitale e diritto alla riservatezza dei dati sanitari nell'era dell'intelligenza artificiale

di *Tommaso Spasari**

SOMMARIO: 1. Introduzione – 2. Il quadro costituzionale e il mutamento degli equilibri tra Stato e Regioni – 3. Il fascicolo sanitario elettronico e l'ecosistema dati sanitari – 4. Il ruolo del Garante per la protezione dei dati personali – 5. La tutela rafforzata dei dati sanitari. – 6. L'uso secondario dei dati per la ricerca scientifica – 7. Lo *jus superveniens* e il bilanciamento tra innovazione e tutela dei diritti – 8. Il consenso informato e la problematica dell'opacità algoritmica – 9. La nuova tripartizione della relazione di cura: medico, paziente, macchina – 10. Osservazioni conclusive.

1. Introduzione

Il progressivo avanzamento tecnologico dei sistemi di intelligenza artificiale (IA) ha profondamente inciso su ogni ambito dell'attività umana, trovando nel settore sanitario un terreno particolarmente fertile per lo sviluppo di applicazioni[1] di *e-health* che promettono di rivoluzionare la prevenzione, la diagnosi e la terapia. Questo fenomeno è stato definito da alcuni come la quarta rivoluzione industriale; inoltre è caratterizzato dall'abilità dell'IA di

* Ph.D.; Professore a contratto di Biodiritto, Università degli Studi Niccolò Cusano.

[1] A. MALVA, V. ZURLO, *Babylon Health: il servizio anglosassone di Medicina Generale privato basato su app,* in «Rivista Società Italiana di Medicina Generale», 5, 2019.

trascendere i confini tralatizi della medicina[2] e del diritto, sollevando questioni deontologiche[3] e biogiuridiche di primaria importanza per la tutela dei diritti fondamentali. L'impiego delle risorse digitali nella sanità non solo potenzia e migliora le prestazioni, ma permette anche di estendere le cure a fasce di popolazione più svantaggiate o residenti in luoghi remoti.

Infatti, l'IA medica offre capacità elaborative e predittive straordinarie, in grado di analizzare e comparare rapidamente una mole ingente di dati per formulare ipotesi diagnostiche o terapeutiche, come la previsione di infarti o l'identificazione di alterazioni del linguaggio del paziente correlate a neuropatie. Il campo d'azione di questi strumenti tecnologici si estende dalla telemedicina[4] e dai dispositivi indossabili[5], per il monitoraggio costante delle funzioni vitali, fino alla chirurgia robotica[6] e alle terapie digitali (DTx). Perciò, questo progresso si traduce in un potenziale aumento dell'accuratezza diagnostica e in una crescente "personalizzazione" delle cure.

Tuttavia, l'implementazione su vasta scala di queste tecnologie porta con sé notevoli "lati oscuri" e problematiche giuridiche, in quanto il diritto si trova spesso a inseguire uno sviluppo tecnologi-

[2] F. GIBELLI, G. RICCI, *L'intelligenza artificiale generativa al servizio del consulente tecnico e del perito: opportunità e sfide,* in «Riv. It. Med. Leg.», 3-4, 2024, 691.

[3] A. PANTI, *La deontologia all'alba dell'intelligenza artificiale,* in «Il punto - confronti su medicina e sanità», 2022.

[4] MINISTERO DELLA SALUTE, *Linee di indirizzo per i servizi di telemedicina*, in G. U. Serie Generale n. 298, 22/12/2022.

[5] S. MAJUMDER et al., *Wearable Sensors for Remote Health Monitoring,* in «Sensors», 1, 2017.

[6] G.Z. YANG *et al.*, *Medical robotics – Regulatory, ethical, and legal considerations for increasing levels of autonomy,* in «Science Robotics», 2017.

co in costante movimento[7]. Perciò, la sfida regolatoria, sia a livello europeo che nazionale, si è concentrata sulla necessità di garantire che l'IA rimanga antropocentrica e affidabile.

Alla luce di quanto sopra prospettato, il processo di digitalizzazione del sistema sanitario nazionale (SSN) e dei sistemi regionali si fonda su pilastri tecnologici essenziali, tra cui spiccano l'interoperabilità dei sistemi informativi, l'analisi delle ingenti quantità di dati conservati e la necessaria interconnessione tra le diverse banche dati. Peraltro, contrariamente a una percezione comune, la normativa sulla protezione dei dati personali non si pone come un ostacolo insormontabile all'innovazione e alla digitalizzazione del SSN. Infatti, essa agisce piuttosto come un motore propulsore, indirizzando in maniera corretta gli interventi legislativi e permettendo una ricognizione precisa dei flussi informativi, per assicurare uno sviluppo equilibrato e costituzionalmente sostenibile dei servizi sanitari erogati alla cittadinanza.

La necessità della digitalizzazione standardizzata e accentrata dei dati sanitari ha ottenuto una rilevanza prioritaria e ineludibile nella riflessione giuridica, specialmente a seguito della crisi pandemica e in vista dei recenti sviluppi regolatori riguardanti l'ecosistema dei dati sanitari (EDS) e la telemedicina. Infatti, l'emergenza sanitaria globale ha costituito un banco di prova inedito che ha evidenziato le frizioni tra i diversi livelli di *governance*, rendendo manifesta l'esigenza di un riassetto delle competenze istituzionali.

Pertanto, l'esperienza pandemica[8] ha spinto verso un rafforza-

[7] A. BIANCARDO, *Problematiche etico-giuridiche relative all'utilizzo dell'intelligenza artificiale in ambito sanitario*, in «Jus-online», 3, 2021.

[8] G. PONZANELLI, *Responsabilità civile, pandemia e sostenibilità*, in «Danno e responsabilità», 2022, 7.

mento delle funzioni centrali dello Stato, mirando a un coordinamento più efficace, soprattutto nel contesto della digitalizzazione. Questa tendenza centripeta è stata ulteriormente accentuata dal piano nazionale di ripresa e resilienza (PNRR)[9], che ha individuato nella trasformazione digitale uno degli assi fondamentali per il rilancio del settore sanitario, garantendo un afflusso consistente di risorse tecniche e finanziarie. Questo impulso ha consolidato il ruolo del governo centrale, ma ha anche accelerato in modo significativo il processo di attuazione del fascicolo sanitario elettronico (FSE)[10].

La digitalizzazione del settore sanitario ha portato in primo piano la questione di natura specificamente costituzionale[11] del delicato e talvolta fragile bilanciamento tra il diritto individuale alla salute, sancito dall'art. 32 Cost. come "diritto fondamentale", e l'interesse collettivo correlato alla tutela della salute pubblica.

2. Il quadro costituzionale e il mutamento degli equilibri tra Stato e Regioni

La nostra Costituzione prevede *ex professo* la tutela della salute, riconoscendone la duplice natura di diritto del singolo e di interesse della collettività. Questa impostazione del quadro costituzio-

[9] P. ZUDDAS, *Intelligenza artificiale in medicina: alcune risposte – significative, ma parziali – offerte dal codice di deontologia medica,* in «Rivista italiana di informatica e diritto», 2, 2024, 579.

[10] C. FILIPPI, *L'impatto della digitalizzazione del SSN sulla protezione dei dati personali*, in G.C. FERONI (a cura di), *Le nuove frontiere della medicina*, Il Mulino, Bologna, 2024, 139.

[11] C. CASONATO, *Intelligenza artificiale e diritto costituzionale: prime considerazioni*, in «Diritto pubblico comparato ed europeo», n.s., 2019.

nale di riferimento si intreccia con il riparto tra i diversi livelli di competenze statali e regionali in questo specifico ambito, che ha visto una graduale ridefinizione dei rapporti tra lo Stato centrale e le Regioni. Infatti, l'art. 117 Cost. attribuiva storicamente l'assistenza sanitaria e ospedaliera alla competenza legislativa concorrente delle Regioni a statuto ordinario, affiancata dal principio di parallelismo sancito dall'art. 118 Cost. che collegava le competenze legislative a quelle amministrative.

Tuttavia, la Riforma del Titolo V della Costituzione[12], risalente al 2001, ha cambiato il riferimento all'assistenza sanitaria con il concetto più ampio di "tutela della salute"[13], pur mantenendola nella competenza concorrente[14]. Nonostante parte della dottrina abbia ipotizzato uno spostamento verso la competenza residuale regionale, la giurisprudenza costituzionale ha generalmente confermato l'organizzazione sanitaria all'interno della competenza concorrente tra Stato e Regioni[15]. Altresì, l'art. 118 Cost. ha abbandonato il principio del parallelismo in favore di quello di sussi-

[12] C. BOTTARI, *La tutela della salute: lavori in corso*, Giappichelli, Torino, 2020, 15.

[13] D. MORANA, *La salute come diritto costituzionale*, Giappichelli, Torino, 2021.

[14] G. CARPANI, D. MORANA, *Le competenze legislative in materia di «tutela della salute»*, in R. BALDUZZI e G. CARPANI (a cura di), *Manuale di diritto sanitario*, il Mulino, Bologna, 2013, 114; M. RUOTOLO, *Il riparto delle competenze legislative tra Stato e regioni, a dieci anni dalla riforma costituzionale,* in «Diritto e società», 2011, 114; F. BENELLI, *Recenti tendenze della Giurisprudenza costituzionale sul riparto per materie tra Stato e regioni e sul declino del principio di leale collaborazione*, in «Le Regioni», 2012, 623; A. D'ATENA, *Diritto regionale*, III ed., Giappichelli, Torino, 2017, 146.

[15] E. GALLO, *La concorrenza nell'erogazione dei servizi sanitari e la posizione delle imprese private*, in «Sanità pubblica e privata», 3, 2003, 256.

diarietà, assegnando le funzioni amministrative ai Comuni, salvo l'esigenza di allocazione a livelli superiori.

Peraltro, nel contesto della sanità digitale assume un ruolo cruciale l'art. 117, comma 2°, lett. m) Cost., poiché riserva allo Stato la competenza esclusiva nella definizione dei livelli essenziali delle prestazioni (LEP) garantite su tutto il territorio nazionale. Inoltre, è molto importante anche la previsione della successiva lett. r) del succitato 2° comma che assegna allo Stato il «coordinamento informativo statistico e informatico» tra le pubbliche amministrazioni. Difatti, questa esplicita disposizione costituzionale rappresenta il referente normativo *in subiecta materia* che legittima il crescente accentramento dei processi decisionali e operativi nel settore della sanità digitale[16].

A fortiori, la centralità dei dati sanitari è ormai indispensabile per l'intero processo assistenziale, purtuttavia esige un sistema di digitalizzazione avanzata e un'efficace interconnessione delle banche dati, con l'obiettivo di garantire l'interoperabilità e la continuità informativa. Pertanto, questa spinta tecnologica ha comportato un rafforzamento del ruolo centrale dello Stato nel coordinamento e nel controllo dei dati sanitari, incidendo così sui tralatizi equilibri costituzionali.

3. Il fascicolo sanitario elettronico e l'ecosistema dati sanitari

Il fascicolo sanitario elettronico (FSE), introdotto con il d.l. n. 179/2012, è stato predisposto come lo strumento fondamentale

[16] G. FIORIGLIO, *Trattamento dei dati sanitari e ricerca medica: profili informatico-giuridici e istituzionali nel contesto europeo*, in «Jura gentium – online first», 1.

per governare la sanità digitale italiana. Il FSE è strutturato come un sistema informativo destinato a tracciare la storia medica dei pazienti, supportando finalità che spaziano dalla diagnosi, alla terapia, alla ricerca, fino alla programmazione, valutazione e controllo dell'assistenza sanitaria. L'istituzione del FSE[17] ha segnato l'introduzione di un modello di *governance* sanitaria fortemente centralizzato, fondato sulla gestione unificata e coordinata della sanità digitale; però, questo approccio centripeto ha comportato un'inevitabile riduzione dell'autonomia decisionale delle Regioni.

Difatti, l'evoluzione normativa del FSE ha subìto una trasformazione radicale, passando da un sistema basato sulla volontarietà del singolo paziente a un modello in cui l'utilizzo del fascicolo è divenuto obbligatorio e automatico. Al riguardo, il d.l. n. 34/2020 ha abrogato la disposizione che condizionava l'acquisizione dei dati sanitari nel FSE al consenso informato[18] dell'interessato, stabilendone l'inserimento automatico[19], inclusi quelli prodotti da operatori privati accreditati, entro cinque giorni dalla prestazione *de qua*. Questa modificazione normativa riflette un evidente spostamento della posizione del legislatore verso la prevalenza dell'interesse collettivo sulla libertà individuale, in quanto viene

[17] V., *amplius*, G.C. FERONI, *Sanità digitale, tra innovazione e tutela dei dati sanitari: il ruolo del garante privacy*, in «Corti supreme e salute», 1, 2025, 225; A. CINQUE, *Privacy, big data e contact tracing: il delicato equilibrio tra diritto alla riservatezza ed esigenze di tutela della salute*, in «La Nuova giurisprudenza civile commentata», 2021, 957.

[18] A. DI MARTINO, *Il consenso trasparente* (*e non solo informato*) *nell'età degli algoritmi: qualche chiarimento dalla Cassazione*, in «Osservatorio sullo Stato digitale IRPA», 2024.

[19] D. AMRAM, *Algoritmi, danno alla persona e nuove soluzioni legal tech,* in «Ciberspazio e diritto», 2021.

limitata *ope legis* – in modo significativo – l'autodeterminazione del cittadino nella gestione dei propri dati sanitari.

Il potenziamento del FSE 2.0 rappresenta un pilastro della Missione 6 "Salute" del PNRR, giacché è strettamente interconnesso con la realizzazione dell'ecosistema dei dati sanitari (EDS). Infatti, l'EDS è uno strumento strategico concepito per raccogliere ed elaborare i dati trasmessi dalle strutture sanitarie e socio-sanitarie e dagli enti del SSN, al fine di garantire il coordinamento informatico e assicurare servizi omogenei su tutto il territorio nazionale. *Ab initio*, lo schema dell'EDS delineava una prospettiva di massiva duplicazione dei dati e dei documenti generati per finalità di cura, creando un vasto *data repository* centrale nazionale dei dati sanitari non pseudonimizzati. Tuttavia, questa configurazione presentava un rischio elevato per i diritti e le libertà dei cittadini interessati, anche in riferimento a possibili rischi di fuga di dati sensibili, giacché implicava trattamenti sistematici su vasta scala di categorie particolari di dati personali *ex* art. 9 GDPR pure mediante logiche algoritmiche[20].

Per questi motivi, il dibattito regolatorio – in particolare a seguito delle osservazioni critiche del Garante per la protezione dei dati personali – ha portato a una profonda rimodulazione dell'EDS. Pertanto, nella sua più recente configurazione l'EDS non rappresenta più un "duplicato" del FSE. Difatti, la nuova struttura prevede un "modulo dati" articolato in componenti distinte e separate per dati in chiaro, dati pseudonimizzati e dati anonimizzati, organizzati in ventuno banche dati regionali più una dedicata ai servizi di assistenza sanitaria al personale navigante (SASN). Sulla scorta di quanto prospettato, attualmente l'EDS ri-

[20] G. RESTA, *Governare l'innovazione tecnologica: decisioni algoritmiche, diritti digitali e principio di uguaglianza,* in «Politica del diritto», 2, 2019.

ceve principalmente i dati che vengono indicizzati nel FSE, inoltre rende disponibili servizi di consultazione e di elaborazione dei dati per finalità di *governance* sanitaria, prevenzione, profilassi e ricerca scientifica, ma solo su richiesta dell'utente, viceversa viene esclusa l'elaborazione automatica e massiva dei dati medesimi.

Altresì, va precisato che le innovazioni tecnologiche del SSN non si limitano al FSE e all'EDS, giacché l'altro pilastro portante della digitalizzazione sanitaria è ravvisabile nell'avvento della c.d. "telemedicina"[21], che integra le recenti tecnologie mediche e informatiche per la gestione a distanza dei percorsi terapeutici.

Nondimeno, l'erogazione di queste prestazioni implica un trattamento massivo di dati sensibili, stabilendo un rigoroso rispetto dei principî di integrità, esattezza e riservatezza e la necessità di una preventiva valutazione d'impatto sulla protezione dei dati (DPIA). Tra l'altro, il Garante ha precisato che la specifica natura di questo trattamento impone comunque massima cautela e trasparenza[22], sebbene per i trattamenti strettamente necessari alla cura – come ad es. la telediagnosi, il teleconsulto o il telemonitoraggio – non sia necessaria l'acquisizione del consenso specifico dell'interessato[23], purché il trattamento sia compiuto soltanto da professionisti sanitari vincolati al segreto professionale.

[21] In dottrina si vedano V. SICA, S. SELVAGGI, M.R. MICHELI, R. BOVA, *Telemedicina. Approccio multidisciplinare alla gestione dei dati sanitari*, Springer, Cham, 2010; E. STEFANINI, *Nuove linee guida nazionali sulla telemedicina: i nodi critici per la piena attuazione,* in «Agenda Digitale», 1° febbraio 2021; R. SENIGAGLIA, *Telemedicina ed essenza fiduciaria del rapporto di cura,* in «Persona e mercato», 3, 2023, 470.

[22] L'art. 50, Reg. UE 2024/1689 prevede gli obblighi di trasparenza inerenti ai sistemi che interagiscono con le persone fisiche.

[23] R. BALDUZZI, D. PARIS, *Corte costituzionale e consenso informato tra diritti fondamentali e ripartizione delle competenze legislative*, in «Giurisprudenza costituzionale», 2008.

Tanto premesso, si evince che l'integrazione della telemedicina con il FSE 2.0 è fondamentale per la continuità assistenziale, ma richiede una solida impostazione di *privacy by design*, tramite l'implementazione di misure di sicurezza adeguate[24] quali la pseudonimizzazione, la cifratura e la resilienza dei sistemi informatici.

4. Il ruolo del Garante per la protezione dei dati personali

Il Garante per la protezione dei dati personali ha giocato un ruolo determinante nella conformazione dell'attuale assetto regolatorio della sanità digitale, spesso intervenendo in via consultiva o suppletiva e colmando le lacune legislative. *Ab initio*, prima della normativa istitutiva del FSE, l'Autorità per la protezione dei dati aveva emanato linee guida che ponevano l'accento sul principio di autodeterminazione e sulla facoltatività della costituzione e alimentazione del FSE, in tal guisa veniva garantito al cittadino il diritto di scegliere i dati da inserire e di manifestare il consenso per la consultazione.

In seguito, dopo il passaggio all'obbligatorietà di alimentazione automatica del FSE post-2020, il Garante *privacy* è intervenuto per mitigare l'impatto dello *jus superveniens* sulla libertà individuale, richiedendo la sussistenza di tre condizioni essenziali per l'inserimento dei dati pregressi che sono ravvisabili nell'implementazione di un'idonea campagna informativa nazionale, nella puntuale comunicazione ai cittadini delle novità introdotte *ope legis*, nonché il riconoscimento di un termine non inferiore a trenta giorni per esercitare il diritto di *opt-out*, cioè manifestare l'eventuale opposizione. Quindi, nel definire la disciplina attuativa del

[24] A. BIANCARDO, *I reati informatici*, Esc Service, Roma, 2018.

FSE 2.0, l'Autorità *de qua* ha ribadito l'esigenza di prevedere manifestazioni di volontà differenziate per ciascuna distinta finalità di trattamento *id est* cura, prevenzione, profilassi internazionale, al fine di garantire l'espressione di un consenso[25] libero, specifico, informato ed esplicito.

Difatti, le criticità di natura normativa e progettuale riscontrate dal Garante *privacy* sono state particolarmente evidenti nello sviluppo iniziale dell'EDS, dove l'Autorità aveva espresso parere negativo a causa di lacune sulla titolarità, la responsabilità, le modalità di alimentazione e, soprattutto, per la massiva duplicazione di dati sanitari in chiaro, contraria al principio di minimizzazione dei dati stabilito *ex* art. 5, par. 1, lett. c) del GDPR. Al riguardo, il Garante ha criticato l'assenza di una visione d'insieme e la mancata considerazione del principio di *privacy by design* fin dalla fase di progettazione[26]. Pertanto, nel successivo schema di EDS è stata espunta la duplicazione del FSE, garantendo così che l'elaborazione dei dati non fosse automatica ma avvenisse soltanto su richiesta e per finalità tassativamente specificate. A questo riguardo, è stato rimosso il riferimento ai sistemi di intelligenza artificiale per l'elaborazione dei dati del FSE 2.0 per l'offerta di servizi EDS, che costituisce un punto cruciale per minimizzare i rischi correlati alle elaborazioni automatizzate nelle prestazioni sanitarie.

Parimenti, in relazione alla telemedicina[27], l'Autorità per la *pri-*

[25] G. GRASSO, *Consenso informato, libertà di scelta e disponibilità del proprio corpo*, in «Forum di Quaderni Costituzionali», 2011.

[26] S. MELCHIONNA, *Sanità digitale e innovazione*, in L. BOLOGNINI, S. ZIPPONI, *Privacy e diritto dei dati sanitari*, Lefebvre Giuffrè, Milano, 2024, 135.

[27] F.E. BROZZETTI, G.M. CANNELLA, A. RANDAZZO, *Telemedicina, teleassistenza e intelligenza artificiale in un sistema socio-sanitario di prossimità: nuovi paradigmi etico-giuridici*, in C. BUZZACCHI, S. LA PORTA (a cura di), *L'integrazione socio-sanitaria e il diritto delle regioni*, Giappichelli, Torino, 2022; P. GUARDA, L.

vacy ha richiesto l'allineamento delle linee guida alla nuova disciplina del FSE e la piena integrazione delle garanzie previste dal GDPR, altresì criticando la carenza di una disciplina organica che assicurasse la coerenza tra i diversi sistemi informativi[28]. Quindi, l'intervento del Garante *de quo* si è dimostrato essenziale per rimodellare il quadro normativo in chiave garantista, limitando i rischi di duplicazione dei dati e rafforzando i principî di *privacy by design* e *by default*.

5. La tutela rafforzata dei dati sanitari

I dati relativi alla salute sono definiti dal GDPR come dati personali attinenti alla condizione fisica o mentale di una persona, inclusa la prestazione di assistenza sanitaria; perciò, rientrano nelle "categorie particolari" il cui trattamento è generalmente vietato *ex* art. 9, par. 1, Reg. (UE) n. 2016/679, salvo le eccezioni previste dal par. 2. Quindi, il divieto è superabile solamente in presenza di specifiche eccezioni, tra cui il consenso esplicito, l'interesse pubblico nel settore della sanità pubblica *ex* lett. i), nonché la ricerca scientifica/statistica *ex* lett. j).

PETRUCCI, *Quando l'intelligenza artificiale parla: assistenti vocali e sanità digitale alla luce del nuovo regolamento generale in materia di protezione dei dati*, in «BioLaw Journal», 2, 2020, 425.

[28] F. LAGIOIA, *L'intelligenza artificiale in sanità: un'analisi giuridica*, Torino, 2020; U. RUFFOLO, *L'Intelligenza artificiale in sanità: dispositivi medici, responsabilità e «potenziamento»*, in «Giurisprudenza italiana», 2, 2021, 502; A. FIORENTINI, *Machine learning e dispositivi medici: riflessioni in materia di responsabilità civile*, in «Il Corriere giuridico», 10, 2021, 1260; G. VOTANO, *Intelligenza artificiale in ambito sanitario: il problema della responsabilità civile*, in «Danno e responsabilità», 6, 2022, 669.

Il dibattito giuridico *in subiecta materia* ha messo in luce una contrapposizione tra il diritto alla protezione dei dati personali e l'impulso alla massimizzazione dell'utilità del dato, accentuata dal fenomeno del "dataismo normativo". In proposito, il dataismo può essere definito come un riduzionismo data-centrico che privilegia la riutilizzabilità e la computabilità a scapito dell'effettiva protezione della dignità umana, giungendo a eclissare alcuni diritti della persona in favore del mero interesse pubblico. Peraltro, il dato è una rappresentazione semplificata e parziale della realtà, quindi il suo utilizzo improprio rischia di fornire una rappresentazione riduttiva dell'individuo, che non può essere considerato una mera somma delle sue parti digitali.

In questo contesto, la Corte europea dei diritti dell'uomo (CEDU) ha rafforzato l'obbligo di tutela, stabilendo che la pubblicazione *online* di decisioni giudiziarie[29] contenenti dati sanitari non anonimizzati impone allo Stato un obbligo positivo di protezione rafforzata. Cosicché, la Corte EDU ha precisato che l'accessibilità perpetua di

[29] S. PELIZZARI, *La protezione dei dati sanitari nelle pubblicazioni giurisdizionali online e gli obblighi positivi dello Stato*, in «Danno e responsabilità», 4, 2025, 472; YU WEITONG, *The Conflict and Balance between Public Disclosure of Judicial Documents and Personal Data Protection: An Examination of the Anonymization Mechanism,* in «Science of Law Journal», 2, 2023, 64–69; F. DONATI, *Trasparenza della giustizia e anonimizzazione dei provvedimenti giudiziari,* in «Astrid Rassegna», 2022; L. LAVRYSEN, *Human Rights,* in *A Positive State: Rethinking the Relationship between Positive and Negative Obligations under the European Convention on Human Rights*, Intersentia, Oxford, 2016; M. E. BONFANTI, *Il diritto alla protezione dei dati personali nel Patto internazionale sui diritti civili e politici e nella Convenzione europea dei diritti umani: similitudini e difformità di contenuti,* in «Diritti umani e diritto internazionale», 3, 2011, 437–481. In giurisprudenza, si v. Sentenza CEDU 13/6/2024, ricorso n. 30782/18; Cass. Civ., Sez. I, n. 10512/2016; sulla raccolta dei dati genetici in ambito penale si v. Corte eur. dir. uomo (Grande Camera), L. B. c. Ungheria, sent. 9/3/2023, ric. n. 36345/16, par. 122.

tali informazioni nell'ecosistema digitale determina una "vulnerabilità permanente" dell'individuo, che è incompatibile con il diritto al rispetto della vita privata *ex* art. 8 CEDU. Ne discende che la protezione dei dati sanitari non si limita a un obbligo negativo di astensione dalle ingerenze, ma impone agli Stati di adottare misure procedurali e normative proattive, come l'anonimizzazione preventiva, che non deve essere vista come un'ec-cezione alla pubblicità bensì come una tecnica di bilanciamento costituzionalmente necessaria per preservare la dignità nell'era digitale. Altresì, la Corte EDU ha finanche definito la pubblicazione online di dati sanitari non anonimizzati[30] come una *particularly serious interference*, vale a dire un'ingerenza di particolare gravità, suscettibile di generare un danno permanente e irreversibile[31], proprio in considerazione della notevole difficoltà di espungere le informazioni sensibili[32] una volta diffuse nei circuiti digitali, c.d. *perpetual availability*.

6. L'uso secondario dei dati per la ricerca scientifica

Il trattamento dei dati sanitari per finalità di ricerca scientifica

[30] P. DATTA *et al.*, *GiusBERTo: Italy's AI-Based Judicial Transformation: A Teaching Case,* in «Communications of the Association for Information Systems», 53, 2023, 751–766.

[31] V., *amplius*, L. G. LOUCAIDES, *Chapter 2. The Concept of "Continuing" Violations of Human Rights*, in *The European Convention on Human Rights*, Brill, Leiden, 2008.

[32] Sul diritto all'oblio sia consentito il rinvio a T. SPASARI, *Profili costituzionali, biogiuridici e medico-legali dell'oblio oncologico*, in «Biodiritto», 1, 2025, 93 ss.; A. GIULIANO, T. SPASARI, *Riflessioni sul diritto alla privacy e la genetica*, in «Medicinaescienzeumane.com», 1, 2025; W. SCHULZ, M. GRAFENSTEIN, *The right to be forgotten*, *in data protection law: a search for the concept of protection,* in «International Journal of Public Law and Policy», 5, 2015, 249.

si articola nella distinzione tra *uso primario* e *uso secondario*. Orbene, l'uso primario si riferisce al trattamento per le finalità iniziali di raccolta come accade nei casi della prestazione di assistenza sanitaria, della diagnosi o della terapia. Viceversa, l'uso secondario riguarda il trattamento per finalità diverse da quelle iniziali, come avviene nell'ambito della ricerca scientifica, della statistica o della *governance* sanitaria.

Al riguardo, è di fondamentale importanza l'art. 110 del D.Lgs. 196/2003, che è stato recentemente novellato *ex lege* 56/2024, in quanto consente di prescindere dal consenso dell'interessato per il trattamento dei dati sanitari acquisiti per fini di ricerca scientifica in campo medico, biomedico o epidemiologico purché la ricerca sia compiuta sulla base di previsioni normative o del diritto dell'Unione Europea, però in conformità all'art. 9, par. 2, lett. j) GDPR. Questa deroga all'acquisizione del consenso *de quo* è consentita anche qualora informare gli interessati risulti impossibile, implichi uno sforzo sproporzionato oppure rischi di compromettere gravemente gli obiettivi della ricerca.

Nelle suddette ipotesi, il titolare del trattamento dei dati è comunque tenuto ad attuare misure appropriate per tutelare i diritti dell'interessato; inoltre, deve ottenere il parere favorevole del Comitato etico competente. Ciononostante, l'introduzione di questa condizione da parte del legislatore, pur eliminando in alcuni casi l'obbligo di consultazione preventiva del Garante della *privacy*, rafforza una tutela fondamentale in quanto affianca il principio di *accountability* del titolare del trattamento con la verifica di un organismo terzo e qualificato.

Inoltre, l'art. 110-bis D.Lgs. 196/2003 regolamenta l'ulteriore trattamento di dati personali da parte di soggetti terzi che svolgano attività di ricerca scientifica o statistica senza il consenso degli interessati, prevedendo in questi casi la necessità di un'autorizzazione del Garan-

te *privacy* qualora l'informativa sia impossibile o sproporzionata. Peraltro, queste modalità di trattamento rendono obbligatoria l'attuazione di misure di anonimizzazione, pseudonimizzazione e minimizzazione dei dati, in conformità a quanto stabilito dall'art. 89 GDPR.

Altresì, nell'ambito di sperimentazioni cliniche – disciplinate dal Regolamento[33] (UE) n. 536/2014 – l'effettuazione degli studi implica il trattamento dei dati relativi alla salute dei partecipanti. Pertanto, in questi casi è necessario il consenso informato alla partecipazione, nonché quello al trattamento dei dati personali per l'uso primario. Inoltre, l'art. 28, par. 2, del Regolamento succitato permette al promotore di richiedere – anche al momento della raccolta del consenso – l'autorizzazione per l'uso secondario dei dati per finalità di ricerca scientifica al di fuori del protocollo, ma il consenso *de quo* resta revocabile. La *ratio legis* sottende il bilanciamento di interessi operato dalla norma tra l'efficacia della ricerca e la protezione dei dati, senza però sacrificare la dignità della persona. Cosicché, un esempio emblematico di questo bilanciamento è ravvisabile nel progetto europeo FACILITATE, che è finalizzato a istituire un ecosistema trasparente tramite il consenso dinamico e un procedimento normativamente conforme per ridurre l'asimmetria informativa, mediante la condivisione e il riutilizzo dei dati sanitari dei partecipanti agli studi clinici.

7. Lo *jus superveniens* e il bilanciamento tra innovazione e tutela dei diritti

L'approccio regolatorio europeo è codificato nel Regolamento

[33] G. RESTA, *Cosa c'è di "europeo" nella Proposta di Regolamento UE sull'intelligenza artificiale?*, in «Dir. inf.», 2022, 323.

UE 2024/1689 (c.d. AI Act[34]) e adotta un modello basato sul rischio, modulando gli obblighi in base al potenziale impatto sui valori giuridici fondamentali. I sistemi di IA utilizzati nell'erogazione di servizi essenziali, tra cui l'assistenza sanitaria, sono classificati come ad alto rischio; perciò, vengono imposti requisiti rigorosi in termini di qualità dei dati, trasparenza e supervisione umana.

Peraltro, il legislatore statale con la legge 23 settembre 2025, n. 132, ha recentemente disciplinato l'impiego dell'intelligenza artificiale (IA) nel settore sanitario, coerentemente con quanto già stabilito dall'AI Act, pur introducendo una disciplina complementare con l'intento di regolare gli aspetti peculiari del contesto italiano. Perciò, questa normativa strategica valorizza la centralità della persona umana, la non discriminazione[35] e la protezione dei dati personali. Inoltre, in particolare nel settore sanitario, l'intervento legislativo *de quo* è finalizzato a promuovere l'impiego dell'IA come supporto nei processi di prevenzione, diagnosi e cura, ribadendo contestualmente che la decisione finale spetta sempre al medico. Infatti, nell'ambito della medicina i sistemi informatici "intelligenti" si stanno progressivamente affermando, pur sollevando complesse questioni giuridiche ed etiche.

In proposito, va precisato che i principî fondamentali per

[34] Il Reg. UE 2024/1689 del Parlamento Europeo e del Consiglio del 13/6/2024, sancisce una disciplina armonizzata sull'intelligenza artificiale e modifica i Reg. CE n. 300/2008; Reg. UE n. 167/ 2013; Reg. UE n. 168/2013; Reg. UE 2018/858; Reg. UE 2018/ 1139 e Reg. UE 2019/2144 e le Dir. 2014/90/UE, Dir. UE 2016/797 e Dir. UE 2020/1828 (regolamento sull'intelligenza artificiale); N. D. GALGANO, *Il Regolamento UE 2024/ 1689 del 13 giugno 2024 sul c.d. alto rischio inaccettabile: le pratiche inerenti sistemi di intelligenza artificiale vietate dal legislatore europeo,* in «Contr. e impr.», 2025, 45.

[35] F. ZUIDERVEEN BORGESIUS, *Discrimination, Artificial Intelligence and Algorithmic Decision-Making*, Council of Europe, 2018.

l'utilizzo dell'IA in sanità sono stabiliti negli articoli 7-10 della legge 132/2025. Pertanto, l'IA può essere impiegata come strumento di supporto nei processi di prevenzione, diagnosi, cura e scelta terapeutica, però senza mai sostituire il giudizio clinico umano, giacché la decisione finale è sempre demandata al medico. Infatti, questa previsione mira a prevenire derive tecnocratiche e a riaffermare la centralità dei professionisti sanitari, proteggendoli dal rischio di deresponsabilizzazione dinanzi a decisioni algoritmiche[36].

Peraltro, l'aspetto cruciale per lo sviluppo della sanità digitale riguarda la ricerca, sicché l'art. 8 della legge n. 132/2025 considera di rilevante interesse pubblico i trattamenti di dati, anche personali, eseguiti per la ricerca e la sperimentazione scientifica nella realizzazione di sistemi di IA per finalità di prevenzione, diagnosi, cura, sviluppo di farmaci o tecnologie riabilitative, ecc., purché svolti da soggetti pubblici o privati senza scopo di lucro, o in collaborazione con IRCCS[37]. Quindi, per queste finalità è sempre autorizzato l'uso secondario di dati personali privi di elementi identificativi diretti, senza un ulteriore consenso dell'interessato ove non previsto dalla legge, fatto salvo l'obbligo di informativa generale. Pertanto, il trattamento di dati per anonimizzazione, pseudonimizzazione o sintetizzazione è sempre consentito, previa informativa. Altresì, è prevista anche l'istituzione di una piattaforma di intelligenza artificiale nazionale presso AGENAS per il supporto all'assistenza territoriale, la cui progettazione, realizzazione e titolarità sono attribuite ad AGENAS in qualità di Agenzia nazionale per la sanità digitale. Tuttavia, questa piattaforma – che riceverà solo i

[36] A. SIMONCINI, *L'algoritmo incostituzionale: intelligenza artificiale e futuro delle libertà,* in «Biolaw Journal», 1, 2019.

[37] Si tratta degli Istituti di ricovero e cura a carattere scientifico, di cui al decreto legislativo 16 ottobre 2003, n. 288.

dati strettamente necessari e che erogherà anche i servizi di supporto non vincolanti ai medici e agli utenti – suscita significative perplessità proprio in riferimento alla potenziale responsabilità[38], specialmente per *automation bias*, nonché alla coerenza con le risorse disponibili.

Analogamente, anche il Consiglio d'Europa era già intervenuto con la *Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law* (risalente a maggio 2024) che, pur avendo un carattere meno vincolante dell'AI Act, ha contribuito a definire principî sovranazionali incentrati sulla dignità umana, l'autonomia individuale, la trasparenza[39] e la *reliability* (*rectius*: l'affidabilità) dei sistemi AI.

Difatti, l'uso di sistemi di IA in medicina solleva interrogativi cruciali sul principio di non discriminazione[40], sancito dall'art. 3 della Costituzione e ribadito dal codice di deontologia medica[41] sia per l'accesso alle cure, sia per la qualità delle prestazioni erogate. Il rischio di operare discriminazioni nell'erogazione delle prestazioni

[38] Cfr., *amplius*, P. CENDON (diretto da), *Responsabilità civile, vol. I*, UTET, Torino, 2020, 3; S. RODOTÀ, *Il problema della responsabilità civile*, in R. NICOLÒ, F. SANTORO-PASSARELLI (diretti da), *Studi di diritto civile*, Giuffrè, Milano, 2023; C. SALVI, *La responsabilità civile*, in G. IUDICA, P. ZATTI (a cura di), *Trattato di diritto privato*, Giuffrè, Milano, 2019, 18; M. BUSSANI, *Le funzioni della responsabilità civile,* in «Riv. dir. civ.», 2022, 264; *Id*., *L'illecito civile*, in P. PERLINGERI (diretto da), *Trattato di diritto civile del Consiglio Nazionale del Notariato*, E.S.I., Napoli, 2020, 113; G. COMANDÈ, *Lettera sulla responsabilità* (*civile*) *e l'autonomia* (*individuale*), in «Danno e responsabilità», 2022, 665.

[39] P. ZUDDAS, *Brevi note sulla trasparenza algoritmica,* in «Amministrazione in cammino», 1, 2020.

[40] D. MORANA, T. BALDUZZI, F. MORGANTI, *La salute "intelligente": eHealth, consenso informato e principio di non-discriminazione*, in «Federalismi.it», 2022.

[41] Il codice di deontologia medica è stato approvato il 18/5/2014 e l'ultima modifica risale al 6/2/2020.

sanitarie è una conseguenza diretta dell'impiego di dati distorti (c.d. *biased*) o incompleti nell'addestramento dei sistemi di *machine learning*.

Peraltro, queste distorsioni generate dai dataset conducono a prestazioni sanitarie differenziate, in quanto i sistemi di IA offrono risposte più adeguate solo ai pazienti appartenenti alle categorie meglio rappresentate. Giova sottolineare finanche un'ulteriore criticità che è rappresentata dalla tendenza dei sistemi di *machine learning* a replicare e ad amplificare le distorsioni apprese, cristallizzando in tal guisa i pregiudizi sociali[42] se non vengono adottati interventi preventivi adeguati, che siano finalizzati a migliorare la selezione dei dati e garantire standard affidabili di test e validazione.

Pertanto, l'art. 7, comma 2°, L. 132/2025 ha introdotto principî di equità algoritmica, stabilendo che l'inserimento di sistemi di IA nel SSN non può in alcun modo selezionare o condizionare l'accesso alle prestazioni sanitarie secondo criteri discriminatori[43], tutelando così il diritto alla salute sancito dall'art. 32 Cost.

8. Il consenso informato e la problematica dell'opacità algoritmica

L'istituto del consenso informato, disciplinato dalla legge n. 219/2017, rappresenta il fondamento normativo della libertà di cura, garantendo il diritto del paziente, sancito *ex professo* dall'art. 32

[42] P. ZUDDAS, *La pari dignità sociale a tre dimensioni. Posizione, giudizio, trattamento*, Wolters Kluver-Cedam, Padova, 2019.

[43] P. ZUDDAS, *Intelligenza artificiale e discriminazioni*, in AA. VV., *Liber amicorum per Pasquale Costanzo - Diritto costituzionale in trasformazione*, I, *Costituzionalismo, Reti e Intelligenza artificiale*, in «Consulta OnLine», 2020.

Cost., all'autodeterminazione nell'accettazione o nel rifiuto del trattamento sanitario. Orbene, la validità del consenso informato[44] è subordinata alla completezza e alla comprensibilità delle informazioni fornite[45]. Tuttavia, l'introduzione dell'IA in ambito diagnostico e terapeutico pone in crisi questi requisiti fondamentali, in quanto le informazioni fornite al paziente dovrebbero anche includere i dettagli sul funzionamento dei sistemi informatici coinvolti che, invece, non vengono divulgati. Difatti, in questo contesto la completezza informativa implicherebbe che il paziente debba essere informato non solo se sia utilizzato un sistema di IA, ma anche quanto e come tale sistema possa incidere sulla diagnosi o sulla terapia[46]. Perciò, l'obbligo di informare sul ricorso all'IA è stato esplicitamente sancito da organi come il Comitato Nazionale per la Bioetica[47].

Inoltre, anche l'art. 7, comma 3°, L. 132/2025, rafforza il diritto all'informazione del paziente che deve essere reso consapevole dell'impiego di tecnologie di IA nel proprio percorso di cura, giacché l'obbligo deriva anche dai principî che sottendono al consenso informato e consente l'esercizio di un controllo effettivo sul processo decisionale.

[44] Cfr., legge 22 dicembre 2017, n. 219, *Norme in materia di consenso informato e di dichiarazioni anticipate di trattamento*.

[45] Cfr. Cass. Civ., III sez., 20/8/2013, n. 19220; Cass. Civ., III sez., 14/3/2006, n. 5444.

[46] L'art. 33, comma 1°, del codice deontologico sancisce che l'informazione fornita dal medico debba concernere «la prevenzione, il percorso diagnostico, la diagnosi, la prognosi, la terapia, le eventuali alternative diagnostico-terapeutiche, i prevedibili rischi e complicanze, i comportamenti che il paziente dovrà osservare nel percorso di cura».

[47] V., COMITATO NAZIONALE PER LA BIOETICA, *Intelligenza artificiale e medicina: aspetti etici*, 2020, 14; nonché CONSIGLIO SUPERIORE DI SANITÀ, *I sistemi di intelligenza artificiale come strumento di supporto alla diagnostic*a, 2021, 30.

Nondimeno, il vero nodo problematico risiede nella necessità di illustrare la logica utilizzata dal sistema. Questa esigenza si scontra con la "doppia opacità" dei sistemi di *machine learning*, ovverosia l'opacità "commerciale" inerente ai segreti industriali che proteggono gli algoritmi, nonché l'opacità "intrinseca" o *black-box*,[48] vale a dire i processi decisionali che sono basati su correlazioni statistiche complesse e spesso non sono neanche note agli stessi programmatori che gestiscono i sistemi di IA. Per questi motivi, l'Organizzazione Mondiale della Sanità ha rilevato che sovente manca una piena comprensione di come tali sistemi di AI possano funzionare[49].

Peraltro, l'obbligo informativo dovrebbe estendersi ai rischi prevedibili derivanti dall'uso di AI, come ad esempio i *bias* algoritmici o gli attacchi informatici. Al riguardo, l'art. 7, comma 6°, L. 132/2025 impone requisiti stringenti di sicurezza e affidabilità, prevedendo che i sistemi di IA e i dati impiegati debbano essere periodicamente verificati e aggiornati per minimizzare il rischio di errori e migliorare la sicurezza dei pazienti. Questa logica di *safety by design* responsabilizza in modo differenziato i fornitori e gli utilizzatori, richiedendo adeguate misure di *cybersecurity*. Perciò, in caso di danno al paziente, si profila un regime di responsabilità potenzialmente assimilabile a quello dei prodotti difettosi di cui alla direttiva 2853/2024/UE[50].

[48] J. M. FELDMAN et al., *Black Box, Gray Box, Clear Box*? *How Well Must We Understand Monitoring Devices*?, in «Anesthesia & Analgesia», 6, 2012.

[49] WORLD HEALTH ORGANIZATION, *Regulatory considerations on artificial intelligence for health*, by Dr. J. Farrar, Chief Scientist, WHO, 2023.

[50] La nuova direttiva per danno da prodotti difettosi è la n. 2853/2024 che abroga la Dir. 85/374/CEE del Consiglio; Cfr. G. PONZANELLI, *Responsabilità da prodotto difettoso*: *dalla Dir. N. 374 del 25 luglio 1985 alla nuova Dir. N. 2853 del 23 ottobre 2024,* in «Danno e responsabilità», 2025, 153; M. FRANZONI, *Note sulla*

A fortiori, l'impiego delle tecniche di *Explainable AI* (XAI) si è dimostrato fondamentale per identificare i fattori che hanno influenzato l'impostazione algoritmica, garantendo un certo grado di interpretabilità. In proposito, la giurisprudenza – nel contesto del trattamento dei dati – ha già richiesto che l'interessato conosca l'*iter* logico sotteso all'*output* per un consenso valido, estendendo in tal guisa i requisiti di trasparenza.

9. La nuova tripartizione della relazione di cura: medico, paziente, macchina

L'introduzione dell'IA in medicina modifica *ab imis* la relazione medico-paziente, trasformandola progressivamente dall'originaria alleanza terapeutica bilaterale in un triplice rapporto medico-paziente-intelligenza artificiale. A questo riguardo, il codice di deontologia medica[51] stabilisce che la relazione di cura è fondata sull'alleanza terapeutica, la fiducia reciproca e il rispetto dei valori, sottolineando che il tempo dedicato alla comunicazione è tempo di cura. Tuttavia, l'integrazione tecnologica – soprattutto attraverso la telemedicina e i dispositivi di monitoraggio remoto – sebbene sia promettente, se non correttamente effettuata, rischia di condurre a una "disumanizzazione" del rapporto medico-paziente[52]. Infatti, la

nuova direttiva sul danno da prodotti difettosi, in «Danno e responsabilità», 2025, 157; nonché v., l'art. 10, Dir. UE 2024/2853; G. DI MARTINO, *Sulla natura della responsabilità per danno da prodotto difettoso,* in «Danno e responsabilità», 2023, 363; Cass., 29/5/2013, n. 13458.

[51] V., *amplius*, art. 20 Codice deontologico.

[52] A. KERASIDOU, *Artificial intelligence and the ongoing need for empathy, compassion and trust in healthcare*, in «Bulletin of the World Health Organization», 1, 2020.

medicina è un'arte fondata su una scienza, ma è composita e antropologica[53], quindi la valutazione clinica basata soltanto su rappresentazioni di dati quantitativi rischia di trascurare la dimensione umana, sociale ed emotiva del paziente[54], compromettendo la percezione olistica del suo benessere[55], ovverosia quello che viene anche definito il *dilution of the patient's account of well-being*. In un simile contesto, il ruolo del medico verrebbe svalutato in quanto risulterebbe degradato ad un mero "intermediario" o una sorta di "traduttore" tra l'*output* del sistema IA e l'assistito medesimo. Alla luce di quanto sopra prospettato – le recenti disposizioni legislative, tra cui la L. 132/2025, intervengono in questo delicato contesto ribadendo la funzione di supporto dell'IA, nonché lasciando impregiudicate le facoltà decisionali finali che vengono demandate dallo *jus superveniens* al professionista sanitario. Va precisato che questo principio dello *human in command* è fondamentale per evitare i fenomeni di *automation bias*, *id est* l'eccessivo affidamento sulle "capacità" della macchina, nonché il c.d. *deskilling*, vale a dire la dequalificazione professionale dei medici per la perdita di capacità critica.

Inoltre, la crescente complessità tecnica dei dispositivi intelligenti e l'opacità dei loro processi decisionali rendono difficile l'individuazione della responsabilità per il danno, ovvero se questo sia attribuibile all'errore umano o a un difetto del macchinario. Quindi, il sistema delineato dall'UE risponde a questa incertezza

[53] In argomento, sia consentito rinviare a T. SPASARI, *Il valore funzionale delle linee guida in materia sanitaria*, in «BioLaw Journal», 1, 2025, 434 ss.

[54] B. MITTELSTADT, *The impact of Artificial Intelligence on the doctor-patient relationship*, Council of Europe, 2021.

[55] L. PERFRANCESCHI, L. PROCURANTI, *Ascolto ed empatia: riflessioni sulla costruzione della relazione di cura,* in A. MORETTO (a cura di), *Il paziente tra coscienza e relazionalità*, Cortina, Verona, 2012.

attraverso il doppio binario della prevenzione mediante l'AI Act e i rimedi risarcitori contemplati dalla nuova direttiva sulla responsabilità da prodotto difettoso[56]. Invece, nel settore sanitario la responsabilità del medico rimane ancorata al consolidato modello della dicotomia tra la responsabilità aquiliana *ex* art. 2043 c.c. e quella da inadempimento contrattuale *ex* art. 1218 c.c., richiedendo una diligenza qualificata nell'interazione con la macchina.

Viceversa, la responsabilità oggettiva[57] potrebbe trovare applicazione nei confronti del fabbricante del dispositivo medico qualora ci fosse un difetto tecnico e, talvolta, anche a carico della struttura sanitaria a causa delle eventuali carenze organizzative, nonché a norma dell'art. 1228 c.c. Al riguardo, la direttiva (UE) 2024/2853 sulla responsabilità per danno da prodotti difettosi[58] introduce importanti aggiornamenti normativi[59], estendendo la nozione di "prodotto" anche al software e ai file per la fabbricazione digitale.

[56] E. BELLISARIO, *Il pacchetto europeo sulla responsabilità per danni da prodotti, anche causati dall'IA. Prime riflessioni sulle proposte della Commissione,* in «Danno e responsabilità», 2023, 153; F. CAROCCIA, *Ancora su responsabilità civile e uso delle intelligenze artificiali,* in «Contr. e impr.», 2022, 408; C. LEANZA, *Intelligenza artificiale e diritto: ipotesi di responsabilità civile nel terzo millennio,* in «Resp. civ. prev.», 2021, 1011; L. ARNAUDO, R. PARDOLESI, *Ecce robot. Sulla responsabilità dei sistemi adulti di intelligenza artificiale,* in «Riv. dir. comm.», I, 1992, 409.

[57] C. CASTRONOVO, *La nuova responsabilità civile*, Giuffrè, Milano, 2006, 275; M. TAMPIERI, *L'intelligenza artificiale e le sue evoluzioni*, Wolters Kluwer-CEDAM, Milano–Padova, 2022, 72; F. DI CIOMMO, *Evoluzione tecnologica e regole di responsabilità civile*, E.S.I., Napoli, 2003, 161.

[58] E. BELLISARIO, *Il pacchetto europeo sulla responsabilità per danni da prodotti, anche causati dall'IA. Prime riflessioni sulle proposte della Commissione,* in «Danno e responsabilità», 2023, 153.

[59] E. PAVAN, *Intelligenza artificiale e modelli di responsabilità civile: riflessioni a margine dei recenti interventi normativi europei,* in «Danno e responsabilità», 6, 2025, 678.

Quindi, la direttiva *de qua* alleggerisce l'onere probatorio per il danneggiato introducendo presunzioni di "difettosità" e di nesso causale, in particolare nei casi in cui il difetto tecnico si manifesti a seguito dell'apprendimento post-immissione sul mercato o per l'inosservanza degli obblighi di aggiornamento di sicurezza.

Tuttavia, l'incertezza sul regime della responsabilità civile per i danni causati dall'IA in sanità persiste[60], in particolare dopo il ritiro per mancanza di consenso tra gli Stati membri dell'*AI Liability Directive* – AILD[61], cioè la proposta di direttiva europea sulla responsabilità extracontrattuale da IA. Pertanto, nel nostro ordinamento andranno definiti univoci criteri di imputazione del danno da IA e modalità probatorie certe, soprattutto nei casi di opacità algoritmica.

10. Osservazioni conclusive

L'attuale evoluzione normativa della sanità digitale delinea un ecosistema complesso; infatti, la transizione verso modelli di *governance* basati sull'accentramento informativo e sul trattamento massivo di dati sanitari solleva molteplici criticità. Il consolidamento del fascicolo sanitario elettronico (FSE) e dell'ecosistema dati sanitari (EDS) quali pilastri della programmazione sanitaria riflette un orientamento regolatorio improntato sulla preminenza dell'interesse pubblico, sebbene finalizzato a scopi costituzionalmente rilevanti come la prevenzione, la ricerca e la profilassi. In questo

[60] G. COMANDÈ, *La r.c. dopo la L. n. 132/2025: per favore parliamone*!, in «Danno e responsabilità», 6/2025, 673.

[61] Il ritiro è stato pubblicato nella Gazzetta Ufficiale dell'Unione Europea (C/2025/5423).

scenario, l'attività del Garante per la protezione dei dati personali si è rivelata determinante per integrare i paradigmi di *privacy by design* e *by default*, salvaguardando l'autodeterminazione informativa mediante il diritto di opposizione all'alimentazione automatica dei flussi.

L'introduzione della Legge 132/2025 in materia di intelligenza artificiale ribadisce la natura antropocentrica del nostro ordinamento, configurando l'IA come mero supporto decisionale ma demandando al medico la decisione clinica. Tuttavia, l'integrazione sistemica tra telemedicina, FSE e algoritmi predittivi impone nuove riflessioni in riferimento alla trasparenza algoritmica e alla minimizzazione del trattamento su vasta scala, rendendo necessari protocolli rigorosi di cybersicurezza e resilienza delle infrastrutture critiche.

In definitiva, affinché la digitalizzazione del servizio sanitario nazionale risulti compatibile con l'assetto dei valori costituzionali è indispensabile che l'innovazione tecnologica non diventi un fine autonomo, ma resti subordinata alla tutela dei diritti fondamentali. Pertanto, la conformità dei futuri provvedimenti attuativi dovrà garantire che la tutela della dignità individuale sia sempre di primaria importanza, impedendo che l'efficienza gestionale pregiudichi la libertà personale e l'equità d'accesso alle cure.

Pertanto, la vera sfida in prospettiva futura non riguarda solamente l'individuazione del criterio più opportuno per l'imputazione della responsabilità[62] e neanche la garanzia della mera traspa-

[62] G. PONZANELLI, *Sostenibilità delle regole di responsabilità civile,* in «Contr. e impr.», 2023, 716; cfr. *Id*., *La irrilevanza costituzionale del principio di integrale riparazione del danno*, in AA. VV., *Giustizia costituzionale e responsabilità civile*, E.S.I., Napoli, 2006. Sul punto v. pure la sentenza Corte cost. 16/10/2014, n. 235, nella quale la Consulta si è pronunciata sulla legittimità costituzionale dell'art. 139, D.Lgs. n. 209/2005, in riferimento al meccanismo tabellare di risarcimento

renza algoritmica. Infatti, a mio avviso, l'autentica *quidditas* risiede nel recupero della dimensione antropologica bilaterale del rapporto di cura medico-paziente[63]. Perciò, è essenziale che l'innovazione tecnologica non determini un'apatia tecnologica o un neo-paternalismo algoritmico, ma sia sempre subordinata al rispetto della persona umana in sintonia con i principî sanciti dalla Carta costituzionale. Quindi, a mio modo di vedere, la relazione tra medico e paziente deve rimanere un'interazione tra due persone, un luogo di incontro e di fiducia che valorizzi l'umanità di entrambe le parti, mentre l'IA può divenire un utilissimo strumento[64] di supporto ma certamente non un sostituto per il *medicus humanus*. In definitiva, il successo dell'IA in medicina dipenderà molto anche dalla capacità del diritto[65] di garantire che la macchina sia sempre al servizio dell'uomo e non viceversa.

del danno biologico introdotto dalla norma *de qua*, nonché Cass. Civ., 11/11/2008, n. 26972.

[63] F. FERRETTI, *La relazione di cura e il consenso informato nell'era della Medical Artificial Intelligence*, in «Corti supreme e salute», 1, 2025, 99.

[64] C. VADALÀ, *L'uso dell'intelligenza artificiale in ambito sanitario: riflessioni a partire da una sperimentazione per lo sviluppo di un SAMD per la diagnosi di autismo infantile,* in «BioLaw Journal», s.i. 1, 2024, 375.

[65] C. NARDOCCI, *La (seconda) svolta del 2024. Anche il Consiglio d'Europa decide di regolamentare l'intelligenza artificiale*, in «BioLaw Journal», s.i. 1, 2024, 73.

L'"algo-libertà": i prodromi del "diritto all'innovazione"

di *Alessandro Sterpa*[*]

SOMMARIO: 1. Lunghezza e profondità delle innovazioni sulla relazione potere-libertà – 2 . L'"algo-libertà" come espansione della sfera dei diritti – 3. Ripensare la forma di Stato attraverso il "diritto all'innovazione".

1. Lunghezza e profondità delle innovazioni sulla relazione potere-libertà

Immergere il diritto nella evoluzione tecnologica è opera davvero complessa per molte ragioni. Prima di tutto, perché quando lo riporti fuori devi fare i conti con il fatto che è cambiata del tutto l'idea che ne avevi. È cambiato anche perché ne sono meglio individuati i tratti di debolezza, proprio come Achille e quel tallone che Teti non riuscì ad immergere nello Stige, senza però che nel frattempo ne sia uscito un semi-dio; anzi. I limiti dell'ordine giuridico si fanno più evidenti.

In seconda istanza, così facendo, per il giurista si apre una questione del tutto nuova rispetto al tradizionale perimetro di ricerca: deve comprendere in cosa lo sta immergendo sapendo che gli stu-

* Professore di Diritto costituzionale e pubblico e coordinatore del dottorato di ricerca in Società in mutamento: politiche, diritti e sicurezza, Università degli Studi della Tuscia. Il saggio riprende le riflessioni svolte durante VII Summer School e Conferenza internazionale di Ventotene «Europa e intelligenza artificiale» e pubblicate nel numero speciale della *Rivista di Diritto pubblico europeo Rassegna online*, fascicolo spec. n. 2 del 2026.

diosi delle innovazioni tecnologiche non solo adottano metodi diversi, ma si occupano di una materia – l'evoluzione tecnologica – che sconfina in altre sino a penetrare nella filosofia dell'evoluzione umana e delle sue forme sociali.

Conviene, allora, che il giurista si doti di una serie di punti di riferimento – i più aperti possibile – rigettando ogni ricostruzione dell'evoluzione tecnica e quindi umana[1] di tipo storicistico ossia segnata, come alcuni sostengono, da leggi generali e tendenze che incorniciano i cambiamenti nell'ambito di più ampi disegni, spesso ideali. Preferiamo immergere il diritto nell'idea di una evoluzione della tecnologia che sia al tempo stesso frammentata e mutevole, incapace di disegnare da sola un indirizzo assiologico del cambiamento che, invece, resta il compito della scelta umana. Decidere se il fiume finisca nel mare, nella diga o contro un insediamento umano è una scelta politica e giuridica che non può prescindere da cosa abbia ingrossato il corso d'acqua.

In questo senso, la storia delle innovazioni scientifiche impiega un criterio ordinatore di particolare rilevanza per il nostro campo di indagine; si tratta del concetto di ciclo; esso si riferisce ad una novità che entra – modificandole – nelle dinamiche sociali, si trasforma essa stessa e si assesta con un certo impatto sulla esistenza umana fino, in alcuni casi, ad essere superata da nuovi strumenti o a perdurare nel tempo con caratteri più o meno stabili. I cicli possono essere di lunghezza cronologica e di incidenza variabili ovviamente, quel che conta è che ogni punto del ciclo rappresenta un momento storico nel quale vi è un angolo di visuale diverso sulla innovazione che compie il percorso. Una foto unica ad ogni scatto come pretende la ricostruzione "aperta" ed evoluzionistica a cui giungono Popper e Lorenz da strade diverse: niente storicismo e

[1] Del fatto che la tecnica sia coessenziale all'evoluzione umana ed espressione stessa dell'uomo da ultimo LEONE XIV in *Magnifica humanitas*, 2026.

leggi universali, piuttosto una faticosa osservazione della (o meglio delle) realtà, in una ricerca che, semplicemente, non ha fine[2].

L'uomo è come oggi si presenta sul palcoscenico del mondo in ragione della rivoluzione culturale ossia della sua capacità di astrazione[3], ma soprattutto grazie a ciò che questa capacità ha permesso; *in primis*, la scoperta del fuoco e il suo controllo, vera prima tecnologia che ha impattato in modo radicale sull'evoluzione umana. La potenza di questo impatto è descritta da Daran Acemoglu e Simon Johnson con l'immagine degli scavi archeologici all'interno di una grotta: negli strati più bassi, quelli più risalenti, si trovano le ossa umane frantumate segno del fatto che fossimo cibo per molti animali e in particolare per quelli che nella grotta vivevano. Scavando gli strati più recenti invece apparivano i resti ossei di animali bruciati segno che, grazie al fuoco, gli ominidi erano riusciti a ribaltare la piramide alimentare; insomma, erano riusciti ad invertire i ruoli naturali grazie all'impiego della tecnica[4]. Lo racconta in termini romanzati ma straordinariamente efficaci Roy Lewis nel suo "*Il più grande uomo scimmia del pleistocene*"[5]: l'uomo che intraprese l'azione di gestire il fuoco decise di assumere un rischio, pagando anche il prezzo di errori fino al successo che fu, a quel punto, prima un vantaggio competitivo per una parte e poi un vantaggio per l'intera umanità.

[2] Si rinvia al volume, come prefazione di Dario Antiseri, K. R. POPPER, K. LORENZ, *Il futuro è aperto*, Rusconi, Milano, 1989, ma anche a K. POPPER, *Verso una teoria evoluzionistica della conoscenza*, Armando Editore, Roma, 1994. Sempre a cura di Dario Antiseri è il volume di K. POPPER, *La ricerca non ha fine*, Armando editore, Roma, 2002.

[3] Come ben descritta da Y. HARARI in *Sapiens. Da animali a dèi.*, Bompiani, Milano, 2017.

[4] D. ACEMOGLU, S. JOHNSON, *Potere e progresso. La nostra lotta millenaria per la tecnologia e la prosperità*, Il Saggiatore, Milano, 2023, 40 e ss.

[5] Così tradotto in Italia per i tipi di Adelphi.

L'intelligenza artificiale appare nei suoi impatti molto simile alla gestione del fuoco per una serie di ragioni. In primo luogo, perché saranno necessari molti anni per effettuare un bilancio consuntivo dei suoi effetti sociali, politici e giuridici. Oggi siamo solo all'inizio di un fenomeno nuovo e di un ciclo quindi anche queste riflessioni sono condizionate dal grado di sviluppo e diffusione, oltre che di impiego pratico e di diverso approccio politico-normativo. In seconda battuta, è simile nella misura in cui la profondità dell'innovazione è assolutamente evidente: stanno cambiando le nostre modalità di impiego dei sensi, della mente e del nostro corpo proprio come accaduto con il fuoco, aprendo così lo scenario a nuove formule di azione individuale e sociale proprio come allora quando imparammo a distruggere e a costruire con il fuoco, a definire strumenti nuovi e a mutare finanche l'organizzazione della vita relazionale oltre che l'attività di nutrizione e di occupazione, gestione e difesa degli spazi. Terzo e decisivo aspetto: a decidere della nuova tecnologia e del suo impiego sarà, dal punto di vista regolatorio, la modalità di composizione del *mix* tra decisione pubblica e decisione autonoma dell'individuo. Se fosse esistito lo Stato quando l'uomo ha scoperto il fuoco, sarebbe andata nello stesso modo che abbiamo richiamato? Probabilmente, la pretesa di proteggere l'individuo dai rischi dell'impiego del fuoco avrebbe condizionato l'evoluzione dell'apporto della nuova tecnologia alla vita individuale e sociale. Probabilmente sarebbe andata in modo diverso.

Cosa vediamo da questo punto di osservazione sebbene già domani potrebbe risultare obsoleto? Sintetizzando, possiamo notare che la dimensione corporea dell'uomo è stata di gran lunga stressata dalla tecnologia e, in particolare, lo è il cervello umano che si trova davanti ad una quantità di elementi, dati e stimoli da gestire che – per come è progettato – non riesce a processare come

faceva prima; in precedenza, infatti, accedevamo ai dati e alle relazioni attraverso una certa prossimità fisica che ne permetteva innanzitutto una limitazione quantitativa anche in ragione delle strutture delle nostre comunità.

Oggi invece viviamo in un perenne *overbooking* informativo in ragione della fine dei limiti a produrre e ad accedere a dati grazie al digitale che ha fatto saltare tutti i contenimenti materiali tipici del mondo analogico; limiti, questi, che consentivano un governo dei dati sociali da parte dei media tradizionali[6]. Da qui la grande capacità, messa in evidenza da molti, degli algoritmi e di chi li governa di potersi atteggiare a nuovi mediatori visto che, nel mare delle informazioni, il nostro cervello si adagia sulle proposte di semplificazione (si pensi alle azioni predittive o confermative, alle bolle e ai suggerimenti) e impiega ogni scorciatoia artificiale come una droga che lo affascina: non a caso si è parlato di un "fascino erotico dell'algoritmo" di fronte a così tanti dati[7]. Un potere privato che profila[8] e incide sull'identità umana[9] che si erge sopra l'individuo ma anche sopra gli stati.

A sua volta, il potere pubblico è stato subito visto – accanto al-

[6] Cfr. A. STERPA, *La libertà di comunicazione digitale. Il potere e il costituzionalismo* in M. E. BUCALO, M. CAPORALE E A. STERPA, a cura di, *Diritto pubblico di internet*, Editoriale scientifica, Napoli, 2024, 15 e ss.

[7] Cfr. A. STERPA, *L'ordine giuridico dell'algoritmo: un nuovo ordinamento giuridico* in A. STERPA, a cura di, *L'ordine giuridico dell'algoritmo*, Editoriale scientifica, Napoli, 2024, 9-31; A. STERPA, I. DE VIVO E C. CAPASSO, *L'ordine giuridico dell'algoritmo: la funzione regolatrice del diritto e la funzione ordinatrice dell'algoritmo*, in «Nuovi autoritarismi e democrazie. Diritto, istituzioni, società», n. 2 del 2023, 120-155.

[8] A. GATTI, *Profilazione e diritti fondamentali: modelli a confronto: Europa, Usa, America Latina*, Editoriale Scientifica, Napoli, 2025.

[9] Oltre le notazioni nell'opera citata di Andrea GATTI, cfr. S. TIRIBELLI, *Identità personale e algoritmi*, Carocci, Roma, 2023. Si veda anche A. STERPA, *La difesa della democrazia dalle minacce ibride*, Editoriale Scientifica, Napoli, 2025.

le grandi aziende – come un pericoloso alleato della tecnologia contro l'uomo nella misura in cui avrebbe potuto accedere a molti dati e controllare le persone, con quel capitalismo della sorveglianza che le esperienze delle autocrazie (di cui la Cina rappresenta il modello più sperimentato) ci stanno mostrando.

Gli studiosi, soprattutto i giuristi, si sono concentrati su questo potenziale effetto di controllo che la tecnologia può garantire sugli individui, memori della funzione più importante che il costituzionalismo è stato chiamato a svolgere ossia dividere il potere e garantire i diritti, come è ricordato anche dalla Dichiarazione del 26 agosto 1789 all'art. 16, secondo il quale – per l'appunto – "*ogni società in cui la garanzia dei diritti non è assicurata, né la separazione dei poteri stabilita, non ha una costituzione*". Ma vale anche nello stesso senso la Dichiarazione di indipendenza del 4 luglio 1776 con il diritto alla libertà e alla felicità intesa come autodeterminazione nella dignità umana.

Ma il rapporto tra la tecnologia e il diritto non si può ridurre a questo; non è tutto qui. Esiste un altro modo di approcciare alle nuove tecnologie nella dialettica tra potere pubblico e individuo e lo si può cogliere solamente se si guarda al tema dalla prospettiva della libertà umana e non da quella del potere pubblico. Dobbiamo chiederci: cosa può fare di più e meglio l'essere umano grazie alle nuove tecnologie? Può essere la tecnologia uno strumento di ampliamento soggettivo e oggettivo dei diritti? Per usare le parole, da ultimo, di Leone XIV, quale indirizzo pro-umano può avere?

2. L'"algo-libertà" come espansione della sfera dei diritti

La risposta l'ha fornita la Corte costituzionale, chiamata a pronunciarsi sull'impossibilità per un elettore di sottoscrivere la pre-

sentazione di una lista per le elezioni della Regione Lazio nel 2023. Con la pronuncia n. 3 del 2025, il Giudice delle leggi ha dichiarato l'illegittimità costituzionale degli artt. 9, terzo comma, della legge 17 febbraio 1968, n. 108 "*Norme per la elezione dei Consigli regionali delle Regioni a statuto normale*" e 2, comma 6, del decreto legislativo 7 marzo 2005, n. 82 recante il "*Codice dell'amministrazione digitale*". La ragione è nel fatto che queste norme non prevedono per l'elettore – che non sia in grado di apporre una firma autografa per certificata impossibilità derivante da un grave impedimento fisico o perché si trova nelle condizioni per esercitare il voto domiciliare – "la possibilità di sottoscrivere un documento informatico con firma elettronica qualificata, cui è associato un riferimento temporale validamente opponibile ai terzi".

La decisione si sviluppa lungo un ragionamento che osserva la tecnologia digitale proprio dal lato dei diritti; la normativa, secondo la pronuncia, contrasta con il "principio personalista (art. 2 Cost.), che impone di rilevare che la dignità umana è compromessa ogni volta in cui è lo stesso ordinamento giuridico che trasforma, in forza di un suo divieto o di una sua previsione, in inabile e bisognosa di assistenza una persona che, invece, sarebbe in grado, con propri mezzi, di provvedere a compiere una determinata attività".

Possiamo organizzare i benefici dell'applicazione delle nuove tecnologie in categorie al solo fine di fornire un quadro che ne faciliti il riconoscimento nelle molteplici sedi in cui sono rintracciabili.

La prima categoria riguarda l'accesso ai diritti e, come nel caso trattato dalla citata sentenza della Corte costituzionale, attiene a tutti quei casi in cui il diritto esiste nella sua forma potremmo dire analogica, ma risulterebbe esercitabile da un numero più ampio di soggetti in senso effettivo grazie alla tecnologia. Oltre alla vicenda esaminata, si pensi ad ogni qualvolta la tecnologia sia in grado di abbattere limiti all'esercizio dei diritti sia civili che sociali: istruzio-

ne (*e-learning*), salute (telemedicina), telelavoro e così via. Certo, occorre valutare l'impatto delle tecnologie sulla qualità delle prestazioni erogate, ma in ogni caso si tratta di elementi ovvi che sono fattibili rispetto al quadro normativo in essere.

Facile immaginare come l'impiego della tecnologia possa aiutare anche l'esercizio di altri diritti, primo tra tutti in ragione della digitalizzazione delle informazioni in possesso della PA: per il diritto di accesso attraverso la digitalizzazione dei documenti[10], certo, ma senza dimenticare i vantaggi rispetto al dovere di contribuire alle entrate fiscali pubbliche piuttosto che per le decisioni amministrative più in generale e per quelle giurisdizionali. Dal punto di vista dell'azione del potere pubblico, la tecnologia digitale potrebbe essere uno strumento di potente – laddove possibile – standardizzazione delle decisioni pubbliche che può rendere più forte lo stato di diritto inteso come "prevedibilità" dell'applicazione delle norme.

Un secondo tipo di vantaggi di cui potrebbero beneficiare gli individui sul piano dei diritti dipende dall'impiego da parte di soggetti privati di natura economica delle nuove tecnologie: si pensi all'industria farmaceutica per la definizione dei vaccini o di nuovi medicinali o, ancora, allo sviluppo di presidi medici e macchinari utili agli interventi; o, ancora, si pensi alla logistica e ai trasporti con il conseguente aumento degli strumenti di controllo e verifica della sicurezza intesa come *safety*.

Vi è però un tema: chi decide se si può accedere o meno

[10] Cfr. Adunanza Plenaria del Consiglio di Stato n. 10 del 2 aprile 2020 nella quale si legge "E proprio nella distanza che intercorre tra bisogno di conoscenza e desiderio di conoscenza sta del resto il tratto distintivo che, al di là di ulteriori aspetti, connota l'accesso documentale rispetto a quello civico generalizzato, nel quale la conoscenza si atteggia quale diritto fondamentale (c.d. *right to know*), in sé, che è premessa autonoma e fondamentale per l'esercizio di qualsivoglia altro diritto".

all'impiego della tecnologia? Qual è il confine tra regolazione pubblica dell'impiego (proibizioni e limiti) e autonoma scelta da parte del soggetto (pubblico e privato) che ritiene di ottenere un beneficio in termini di espansione della propria sfera giuridica? Le questioni interrogano il rapporto tra libertà e autonomia della persona (fisica e giuridica) e potere pubblico ossia, per i costituzionalisti, la questione della forma di stato.

3. Ripensare la forma di Stato attraverso il "diritto all'innovazione"

Con il concetto di "forma di stato" intendiamo il rapporto tra libertà e autorità o, se si vuole, tra il soggetto che è sottoposto alle regole pubbliche e chi quelle regole pone. Nella nostra forma di stato liberale e democratica, quel rapporto è definito dall'impianto costituzionale dello stato di diritto (meglio reso dall'espressione *rule of law*) in base al quale, nel cercare di costruire una tendenziale prevedibilità delle conseguenze delle azioni umane – i diritti tra loro si bilanciano per garantire che non esista uno di essi come "tiranno". Nel bilanciamento, si tiene conto dell'impianto assiologico e valoriale della nostra Costituzione in modo che non sia frustrato il principio personalista, ben descritto dalla Corte costituzionale nella citata sentenza, ma che compare più volte nella giurisprudenza del Giudice delle leggi.

Secondo l'art. 2 della Costituzione, "*la Repubblica riconosce e garantisce i diritti inviolabili dell'uomo, sia come singolo, sia nelle formazioni sociali ove si svolge la sua personalità, e richiede l'adempimento dei doveri inderogabili di solidarietà politica, economica e sociale*". In esso convivono il principio personalistico, quello plura-

listico e quello solidaristico[11]. L'indeterminatezza del testo ha spinto la dottrina verso interpretazioni non sempre conformi tra loro, una delle quali, in particolare, ha inteso il binomio "inviolabile" e "riconoscimento" ad un patrimonio di valori e posizioni giuridiche soggettive preesistente allo Stato stesso o, comunque, una "clausola aperta" ad altri valori e altre libertà che possano emergere[12].

Per il principio personalista, quindi, l'individuo è un valore in sé per la comunità e per l'ordinamento giudico, ma anche con una evidenza relazionale per cui, nella sua versione "sociale", essa attiene alla "pari dignità sociale" prevista dall'art. 3 della Costituzione quale protezione del valore della persona nei rapporti umani regolati dal diritto. Vi è quindi, in Costituzione, una abbondanza dei riferimenti alla dignità[13], come limite all'iniziativa economica[14], piuttosto che ai fini della quantificazione della retribuzione del lavoro[15]. Nonostante la propensione, che il testo costituzionale ha

[11] Cfr. B. CARAVITA, *Articolo 2*, in V. CRISAFULLI, L. PALADIN, *Commentario breve alla Costituzione italiana*, Cedam, Padova, 1990, 9 e la *dottrina* ivi richiamata.

[12] C. MORTATI, *Istituzioni di diritto pubblico…* cit., vol. II, 1038; come ricorda B. Caravita, *ult. op. cit*. altri Autori l'hanno ricondotta alla costituzione materiale piuttosto che ai principi e agli orientamenti di fondo dell'ordinamento giuridico. Cfr. anche F. MODUGNO, *I "nuovi diritti" nella giurisprudenza costituzionale…* cit., secondo il quale la libertà come valore comporta che "l'*art. 2* come generale statuizione sul riconoscimento-garanzia dei diritti inviolabili, oltre e indipendentemente dalle specifiche e puntuali previsioni costituzionali dei diritti, può significare in fondo proprio questo: che la libertà come valore non può essere circoscritta a previsioni determinate e specifiche delle esplicazioni e direzioni in cui esso si realizza" (8).

[13] Da ultimo insiste su diversi sensi di "dignità" in rapporto con l'IA Papa Leone XIV nell'Enciclica citata.

[14] L'iniziativa economia privata, secondo l'art. 41 Cost., "*non può svolgersi in contrasto con l'utilità sociale o in modo da recare danno alla sicurezza, alla libertà, alla dignità umana*".

[15] Ai sensi dell'art. 36, comma 1, Cost., "*il lavoratore ha diritto ad una retribu-*

comportato, a declinare la dignità umana in termini di diritti sociali che preservassero l'individuo dalle condizioni di bisogno incompatibili – appunto – con la propria dignità, la giurisprudenza costituzionale l'ha declinata anche in termini di libertà individuale di scelta, di coscienza e di pensiero: "la sfera intima della coscienza individuale deve esser considerata come il riflesso giuridico più profondo dell'idea universale della dignità della persona umana che circonda quei diritti, riflesso giuridico che, nelle sue determinazioni conformi a quell'idea essenziale, esige una tutela equivalente a quella accordata ai menzionati diritti, vale a dire una tutela proporzionata alla priorità assoluta e al carattere fondante ad essi riconosciuti nella scala dei valori espressa dalla Costituzione italiana" (sent. n. 467 del 1991).

La dignità umana, quindi, costituisce un principio fondamentale nella misura in cui consente di liberare il singolo da tutte le condizioni che non gli consentano di esprimere la propria personalità, le proprie inclinazioni e le proprie capacità[16], tanto da trovare collocazione non solo nell'art. 2 della Costituzione (laddove si parla di personalità[17]), ma anche nelle previsioni costituzionali di cui all'art. 4, comma 2, della Costituzione allorché, oltre a riconoscere il diritto al lavoro, esplicita il dovere di tutti di "*svolgere, secondo le pro-*

zione proporzionata alla quantità e qualità del suo lavoro e in ogni caso sufficiente ad assicurare a sé e alla famiglia un'esistenza libera e dignitosa".

[16] Molteplici, in questo senso, i riferimenti in materia di disciplina di pene detentive e di istituzioni carcerarie, incentrata sul "recupero della persona, valore centrale per il nostro sistema penitenziario non solo sotto il profilo della dignità individuale ma anche sotto quello della valorizzazione delle attitudini e delle specifiche capacità lavorative del singolo» (sentenza n. 158 del 2001; nello stesso senso, si è espressa la sentenza n. 341 del 2006).

[17] "*La Repubblica riconosce e garantisce i diritti inviolabili dell'uomo, sia come singolo, sia nelle formazioni sociali ove si svolge la sua personalità, e richiede l'adempimento dei doveri inderogabili di solidarietà politica, economica e sociale*".

prie possibilità e la propria scelta, una attività o una funzione che concorra al progresso materiale e spirituale della società".

A questo punto si evidenzia il grande bivio che incontriamo nel rapporto tra IA e forma di stato: da un lato la via di riflessione che si concentra sul ruolo del potere pubblico come principale soggetto sul quale appostare le esigenze di uno sviluppo dell'IA controllato al fine del rispetto dei principi e dei valori costituzionali, dall'altra la strada che passa per una valorizzazione del ruolo dell'individuo come soggetto in grado di scegliere attraverso la propria autonomia quale ruolo fornire nella sua esistenza all'IA.

Ovviamente non è un bivio che consente di scegliere tra due strade alternative, quanto piuttosto di decidere quale contributo può arrivare da ciascuno dei due approcci che, presi singolarmente, appaiono entrambi inefficaci. Lo Stato non è in grado di "catturare" il fenomeno tecnologico imponendo regole davvero efficaci, come dimostrano i tentativi di regolazione peraltro sottoposti a repentini cambiamenti dell'oggetto da normare. La dimensione sovranazionale, pur maggiormente in grado di incidere, deve comunque operare in un contesto di equilibri globali. L'individuo, da parte sua, è in molti casi condizionato dall'IA anche nella sua sfera cognitiva e, senza un adeguato approccio critico e culturale, ne può ben essere "catturato". In questo contesto, la soluzione appare l'impiego di strumenti statuali ed europei di regolazione, certamente, ma anche la fiducia in un individuo in grado di decidere nel libero mercato della tecnologia cosa sia meglio per lui, superando in certi frangenti anche il dogma della indisponibilità di alcuni diritti.

Questo doppio binario è già evidente nell'ambito della "guerra cognitiva" dove, per l'appunto, la difesa dalle *fake news* e dalle narrazioni condizionanti è effettuata sia con strumenti regolatori che attraverso l'azione selettiva dell'individuo[18].

[18] Sia concesso rinviare a A. STERPA, *La Costituzione e la "nuova guerra": Sta-*

Prevedere la configurazione di un diritto all'innovazione, come per primo ha evidenziato Giovanni Guzzetta, non significa imporlo come un diritto tiranno ma sottoporlo al bilanciamento con gli altri diritti[19].

A chi spetta questo bilanciamento? Ovviamente ai soggetti legittimati eletti dal popolo per adottare le norme a livello nazionale ed europeo, senza dimenticare il ruolo che avrà necessariamente la giurisprudenza italiana e comunitaria.

Quello che conta, a nostro parere, è che il sentiero della centralità dell'individuo non sia scambiato per una negazione del ruolo del potere pubblico quanto, piuttosto, l'approccio che può evitare di cogliere tutte quelle occasioni che si potrebbero permettere di rintracciare nella tecnologia una nuova forma di ampliamento della libertà umana. In questi anni, infatti, si è persa da più parti la prospettiva dell'autonomia personale mentre è accresciuta l'idea della ineluttabilità della presenza di una norma di diritto pubblico. Ogni problema che sorge è sempre più visto come uno spunto di regolazione e non come uno spunto per l'autonomia privata. L'idea sottesa a questo approccio è che l'ordine abbia una progettualità eteronoma rispetto all'individuo, mentre l'individualismo metodologico ci insegna che occorre muovere dall'individuo per costruire un ordine "senza piano" che è molto più efficiente[20].

to e individuo nella guerra cognitiva in *federalismi.it* n. 14 del 2026, 229 ss.

[19] G. GUZZETTA, *Il diritto costituzionale di fronte all'IA* in S. PREZIOSI, a cura di, *Intelligenza artificiale e responsabilità umana*, Tab, Roma, 2026, 67 ss.. Cfr. anche il rapporto della Fondazione Luigi Einaudi del febbraio 2025.

[20] L. INFANTINO, *L'ordine senza piano*, Armando Editore, Roma, 2008; M. FRIEDMAN, *There's No Such Thing As a Free Lunch*, 1975 (*Non esistono pasti gratis*, Liberilibri, Macerata, 2025).

L'uso degli algoritmi nell'analisi del rischio relativa al contrasto agli illeciti fiscali

di *Massimo Varriale*[*]

SOMMARIO**:** 1. Le novità normative – 1.1. Risultati dell'analisi del rischio – 1.2. Unità integrata permanente di analisi del rischio – 2. Rapporto tra analisi del rischio fiscale e intelligenza artificiale – 2.1. Variabili di profilazione

1. Le novità normative

Il decreto legislativo n. 13 del 2024, recante disposizioni in materia di accertamento tributario e di concordato preventivo biennale, all'articolo 2, nel dare attuazione alla delega fiscale, ha operato una "Razionalizzazione e riordino delle disposizioni normative in materia di attività di analisi del rischio"[1] fiscale finalizzate alla pre-

[*] Dottore di ricerca in Diritto Tributario presso l'Università della Tuscia.

[1] "*Il potenziamento dell'attività di analisi del rischio è stato ampiamente valorizzato anche nell'ambito del Piano Nazionale di Ripresa e Resilienza (PNRR), che, a completamento delle* «azioni»*, prevede diverse* «riforme»*, intese quali misure destinate a concorrere alla realizzazione degli obiettivi, già indicati nelle* Country Specific Recommendations *rivolte all'Italia dall'Unione Europea, di equità sociale e miglioramento della competitività del sistema produttivo. Tra queste, quella denominata* «Riduzione del *tax gap*»*, prevede esplicitamente che* «L'obiettivo del potenziamento dei controlli sarà realizzato attraverso selezioni più mirate dei contribuenti a maggiore rischio di evasione, rese possibili dall'applicazione di strumenti di *data analysis* più avanzati e dall'interoperabilità delle banche dati». *Tra gli strumenti di data analysis vengono citate le* «tecniche sempre più avanzate come intelligenza artificiale, *machine learning, text mining,* analisi delle relazioni»" (Relazione illustrativa del Decreto Legislativo 12 febbraio 2024, n. 13, recante: «*Disposizioni in materia di accertamento tributario e di concordato preventivo biennale*», pubblicato nella Gazzetta Ufficiale n. 43 del 21 febbraio 2024).

venzione e al contrasto agli illeciti tributari.

Preliminarmente ha introdotto le definizioni di analisi del rischio, rischio fiscale, criterio selettivo, indicatore di rischio desunto o derivato, analisi deterministica, analisi probabilistica, applicabili "a tutte le disposizioni di legge che richiamano l'analisi del rischio in materia tributaria".

Al riguardo emerge che il Legislatore delegato ha inteso normare le definizioni del processo, non rinviando a disposizioni interne delle diverse pubbliche amministrazioni interessate, creando un linguaggio comune ed intellegibile anche per i relativi stakeholder.

In particolare, viene chiarito che:

a) l'analisi del rischio è "il processo, composto da una o più fasi, che, al fine di massimizzare l'efficacia delle attività di prevenzione e contrasto all'evasione fiscale, alla frode fiscale e all'abuso del diritto in materia tributaria, nonché di quelle volte a stimolare l'adempimento spontaneo, tramite modelli e tecniche di analisi deterministica ovvero probabilistica, nel rispetto della normativa in materia di trattamento di dati personali, utilizza, anche attraverso la loro interconnessione, le informazioni presenti nelle basi dati dell'Amministrazione finanziaria, ovvero pubblicamente disponibili, per associare, coerentemente a uno o più criteri selettivi, ovvero a uno o più indicatori di rischio desunti o derivati, la probabilità di accadimento a un determinato rischio fiscale, effettuando, ove possibile, anche una previsione sulle conseguenze che possono generarsi dal suo determinarsi";

b) il rischio fiscale è il "rischio di operare, colposamente o dolosamente, in violazione di norme di natura tributaria, ovvero in contrasto con i principi o con le finalità dell'ordinamento tributario";

c) un criterio selettivo è dato dall' "identificazione e tipizzazio-

ne di una condotta, monosoggettiva o plurisoggettiva, idonea a concretizzare un rischio fiscale";

d) un indicatore di rischio desunto o derivato è il "risultato di un processo di profilazione finalizzato a ottenere ulteriori caratterizzazioni dei contribuenti oggetto di analisi". Si tratta cioè di un processo che permette di effettuare classificazioni, ovvero attribuire punteggi di rischio;

e) l'analisi deterministica è l'"insieme dei modelli e delle tecniche di analisi basati sul raffronto e sull'elaborazione di dati, riferiti a uno o più contribuenti ovvero a uno o più periodi di imposta, volti a verificare, tramite criteri selettivi fondati su relazioni non probabilistiche, l'avveramento di un rischio fiscale, in tutto o in parte definibile prima dell'avvio dell'analisi";

f) l'analisi probabilistica è l'"insieme dei modelli e delle tecniche di analisi che, sfruttando soluzioni di intelligenza artificiale ovvero di statistica inferenziale, consentono di isolare rischi fiscali, anche non noti a priori, che, una volta individuati, possono essere utilizzati per l'elaborazione di autonomi criteri selettivi, ovvero permettono di attribuire una determinata probabilità di accadimento a un rischio fiscale noto".[2]

Con riferimento al concetto di sistema di intelligenza artificiale l'articolo 3 del Regolamento Europeo 2024/1689 del 13 giugno 2024 (AI Act)[3] dispone che si tratta di "un sistema automatizzato progettato per funzionare con livelli di autonomia variabili e che può presentare adattabilità dopo la diffusione e che, per obiettivi

[2] La statistica inferenziale può essere intesa come il procedimento per cui le caratteristiche di una popolazione si inducono dall'osservazione di una sola parte della stessa; tale parte è selezionata solitamente mediante un esperimento casuale.

[3] L'articolo 113 dell'AI Act con riferimento all' "*Entrata in vigore e applicazione*" dispone che "*entra in vigore il ventesimo giorno successivo alla pubblicazione nella Gazzetta ufficiale dell'Unione europea*" e che "*Si applica a decorrere dal 2 agosto 2026*".

espliciti o impliciti, deduce dall'input che riceve come generare output quali previsioni, contenuti, raccomandazioni o decisioni che possono influenzare ambienti fisici o virtuali".

Il medesimo Regolamento ha previsto che "I sistemi di specificamente destinati a essere utilizzati per procedimenti amministrativi dalle autorità fiscali e doganali, come pure dalle unità di informazione finanziaria che svolgono compiti amministrativi di analisi delle informazioni conformemente al diritto dell'Unione in materia di antiriciclaggio, non dovrebbero essere classificati come sistemi di IA ad alto rischio utilizzati dalle autorità di contrasto a fini di prevenzione, accertamento, indagine e perseguimento di reati".

1.1. Risultati dell'analisi del rischio

Al comma 2 dell'articolo 2 del Dlgs n. 13 del 2024 vengono fornite indicazioni sull'utilizzo dei risultati dell'analisi del rischio che, oltre che per le finalità di prevenzione e contrasto all'evasione fiscale, alla frode fiscale e all'abuso del diritto in materia tributaria, nonché di stimolo dell'adempimento spontaneo, si afferma possano essere utilizzati anche per lo svolgimento di controlli preventivi.

Al comma 3 del medesimo articolo 2 del Dlgs n. 13 del 2024 viene invece disciplinato l'ambito applicativo, con riferimento alle basi dati utilizzabili, dell'analisi del rischio fiscale; al riguardo "Le informazioni presenti in tutte le basi dati di cui l'Agenzia delle entrate dispone[4] […] sono utilizzate dall'Agenzia delle entrate, anche

[4] Ivi comprese quelle presenti nell'apposita sezione dell'Anagrafe tributaria di cui all'articolo 7, sesto comma, del decreto del Presidente della Repubblica 29 settembre 1973, n. 605, e all'articolo 11, comma 2, del decreto-legge 6 dicembre 2011, n. 201, convertito, con modificazioni, dalla legge 22 dicembre 2011, n. 214, nonché quelle memorizzate ai sensi dell'articolo 1, comma 5-bis, del decreto legislativo 5 agosto 2015, n. 127, escluse quelle soggette alla disciplina di cui al decreto legislativo 18 maggio 2018, n. 51.

tramite interconnessione tra loro e con quelle di archivi e registri pubblici per le attività di analisi del rischio fiscale, per le attività di controllo, per le attività di stimolo dell'adempimento spontaneo e per quelle di erogazione di servizi ai contribuenti"[5].

Con riferimento alla tutela degli interessi pubblici sottesi alle attività funzionali alla lotta e all'evasione fiscale, che potrebbero essere lesi dall'esercizio dei diritti degli interessati, il successivo comma 4 dispone che "Limitatamente alle attività di analisi del rischio condotte dall'Agenzia delle entrate, d'intesa con il Dipartimento delle finanze del Ministero dell'economia e delle finanze, nel rispetto delle disposizioni di cui agli articoli 2-undecies, comma 3, del codice in materia di protezione dei dati personali di cui al decreto legislativo" n. 196 del 2003 "nonché dell'articolo 23, paragrafo 1," del GDPR[6] "considerati i principi di necessità e di pro-

[5] L'articolo 9 "*Trattamento di categorie particolari di dati personali*" del GDPR dispone che "1. *È vietato trattare dati personali che rivelino l'origine razziale o etnica, le opinioni politiche, le convinzioni religiose o filosofiche, o l'appartenenza sindacale, nonché trattare dati genetici, dati biometrici intesi a identificare in modo univoco una persona fisica, dati relativi alla salute o alla vita sessuale o all'orientamento sessuale della persona*" eccetto, tra l'altro, quando "*il trattamento è necessario per motivi di interesse pubblico rilevante sulla base del diritto dell'Unione o degli Stati membri, che deve essere proporzionato alla finalità perseguita, rispettare l'essenza del diritto alla protezione dei dati e prevedere misure appropriate e specifiche per tutelare i diritti fondamentali e gli interessi dell'interessato*". In merito si osserva che l'articolo 2 *sexies* del dlgs n. 196 del 2003 ha chiarito che "*si considera rilevante l'interesse pubblico relativo a trattamenti effettuati da soggetti che svolgono compiti di interesse pubblico o connessi all'esercizio di pubblici poteri nelle seguenti materie*:

[…];

i) attività dei soggetti pubblici dirette all'applicazione, anche tramite i loro concessionari, delle disposizioni in materia tributaria e doganale, comprese quelle di prevenzione e contrasto all'evasione fiscale;

[…].

[6] Regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio del 27 aprile 2016.

porzionalità, con regolamento del Ministro dell'economia e delle finanze, sentito il Garante per la protezione dei dati personali, sono definite:

a) le specifiche limitazioni e le modalità di esercizio dei diritti di cui agli articoli 15, 17, 18, 21 e 22 del predetto" GDPR ", in modo da assicurare che tale esercizio non possa arrecare un pregiudizio effettivo e concreto all'obiettivo di interesse pubblico; b) le disposizioni specifiche relative al contenuto minimo essenziale di cui all'articolo 23, paragrafo 2, del predetto" GDPR; c) le misure adeguate a tutela dei diritti e delle libertà degli interessati".

1.2. Unità integrata permanente di analisi del rischio

Il comma 9 dell'articolo 2 del dlgs n. 13 del 2024 in precedenza citato statuisce che "L'Agenzia delle entrate e la Guardia di finanza, per le finalità di prevenzione e contrasto all'evasione fiscale, alla frode fiscale e all'abuso del diritto in materia tributaria, possono, compatibilmente con le vigenti disposizioni in tema di trattamento di dati personali, di riservatezza o di segretezza, mediante la stipula di protocolli d'intesa, condividere tra loro, anche avvalendosi" della SOGEI[7] "tutte le informazioni e le risorse informatiche di cui dispongono, salvo le informazioni soggette alla disciplina di cui al decreto legislativo 18 maggio 2018, n. 51[8], anche tramite la costituzione, conformemente ai rispettivi ordinamenti, di unità integrate

[7] Sogei - Società Generale d'Informatica S.p.A. - è la società di *Information Technology* al 100% del Ministero dell'Economia e delle Finanze; opera sulla base del modello organizzativo dell'*in-house*.

[8] Tale disposizione dà attuazione della direttiva (UE) 2016/680 del Parlamento europeo e del Consiglio, del 27 aprile 2016, relativa alla protezione delle persone fisiche con riguardo al trattamento dei dati personali da parte delle autorità competenti a fini di prevenzione, indagine, accertamento e perseguimento di reati o esecuzione di sanzioni penali, nonché alla libera circolazione di tali dati.

di analisi del rischio"; in tale caso "Al personale destinato a dette unità restano applicabili le disposizioni normative, regolamentari e contrattuali che si applicano alle relative amministrazioni".

Il protocollo di intesa tra l'Agenzia delle entrate e la Guardia di finanza è stato sottoscritto in data 13 marzo 2024 e prevede la costituzione di un'apposita "Unità integrata permanente di analisi del rischio" (UIPAR) finalizzata ad instaurare "forme rafforzate di collaborazione in materia di analisi del rischio per finalità di prevenzione e contrasto all'evasione fiscale".

2. Rapporto tra analisi del rischio fiscale e intelligenza artificiale

Il tema del rapporto tra analisi del rischio fiscale[9] e intelligenza artificiale è stato trattato, da ultimo, nella recente "relazione sull'economia non osservata e sull'evasione fiscale e contributiva anno 2024" predisposta ai sensi dell'articolo 10-bis, comma 3, della legge 31 dicembre 2009, n. 196.

In tale sede viene evidenziato che l'analisi del rischio fiscale è "un'attività che consente di incrementare il livello di conoscenza dell'ambiente esterno, tramite un uso efficiente dei dati a disposizione, così da ridurre il livello di asimmetria informativa in cui si trova ad operare l'Amministrazione finanziaria; l'attività di analisi del rischio fiscale si inserisce, quindi, nell'ambito della gestione del rischio di non conformità (Compliance Risk Management).

In particolare, viene osservato che da circa un decennio, l'OCSE e la Commissione europea hanno fornito indirizzi e condiviso strategie per un'efficace azione di risk management, che risponde all'esigenza di allocare al meglio le risorse delle Ammini-

[9] Cfr. paragrafo precedente.

strazioni fiscali – scarse rispetto alla platea dei contribuenti – al fine di raggiungere un livello di compliance fiscale ottimale.

Si tratta quindi di un processo sistematico in cui l'Amministrazione fiscale effettua scelte di merito in relazione agli strumenti più adeguati a stimolare efficacemente la compliance, ovvero prevenire e contrastare la sua assenza; tale attività è tanto più efficace quanto maggiore è la conoscenza del comportamento dei contribuenti".

La citata relazione osserva che le potenzialità dell'analisi del rischio sono oggi fortemente valorizzate dal rilevante patrimonio informativo[10] di cui l'Amministrazione finanziaria può disporre e di cui potrebbe disporre in prospettiva, in esito anche ad evoluzioni normative ed organizzative, e richiama la rilevanza della sottoscrizione, in data 13 marzo 2024, del Protocollo d'intesa tra l'Agenzia delle entrate e la Guardia di Finanza, in precedenza citato.

In merito, la relazione osserva che l'analisi del rischio per l'individuazione dei fenomeni di evasione, elusione e frode, grazie alla disponibilità e al tempestivo utilizzo delle informazioni rilevanti nella disponibilità dell'Amministrazione finanziaria, consente il raggiungimento di molteplici effetti:

- garantire la prevenzione prima che siano attuati comportamenti evasivi, elusivi o frodatori;

- operare la repressione *ex post*, grazie anche all'attuazione di azioni mirate;

- indurre l'innalzamento complessivo del livello di compliance;

[10] "*Il vasto patrimonio informativo di cui dispone l'Amministrazione finanziaria, per essere correttamente valorizzato, richiede l'utilizzo di strumenti di analisi sempre più evoluti e, conseguentemente, continui investimenti volti ad incrementare le competenze e le conoscenze del personale e ad aggiornare l'infrastruttura tecnologica, così da riuscire a trasformare i dati in informazioni e le informazioni in conoscenza*". (Cit. relazione sull'economia non osservata e sull'evasione fiscale e contributiva anno 2024).

- permettere un uso più efficiente delle risorse dell'Amministrazione, circoscrivendo i controlli nei confronti di soggetti a più alto rischio fiscale e con minore impatto su cittadini e imprese, anche in termini di oneri amministrativi.

In questo modo si dà concreta attuazione ai principi costituzionali che sottendono l'azione dell'Amministrazione fiscale, in primis quelli contenuti agli articoli 53 ("Tutti sono tenuti a concorrere alle spese pubbliche in ragione della loro capacità contributiva. Il sistema tributario è informato a criteri di progressività") e 97 ("Le pubbliche amministrazioni, in coerenza con l'ordinamento dell'Unione europea, assicurano l'equilibrio dei bilanci e la sostenibilità del debito pubblico").

Centrale risulta, quindi, la funzione assunta dalle metodologie di analisi dei dati basate su soluzioni di intelligenza artificiale e, in particolare, sulle tecniche di machine learning, che confluiscono in quello che può essere definito come "approccio probabilistico"[11] [12] all'analisi del rischio fiscale.

L'analisi probabilistica, quindi, consente di attenzionare le fattispecie non associabili "con certezza" a violazioni o irregolarità, che però potrebbero essere meritevoli di analisi interne ed appro-

[11] "*L'approccio probabilistico parte dell'assunto che, in talune circostanze caratterizzate da carenza informativa o cognitiva, la selezione di un contribuente da sottoporre a controllo è un tipico evento aleatorio, che sconta un grado di incertezza relativo all'esito del controllo stesso. Se si riesce ad associare all'evento "controllo" una valutazione di probabilità rispetto all'esito atteso, si trasforma un evento incerto in un evento caratterizzato da un certo grado di rischio*". (Cit. relazione sull'economia non osservata e sull'evasione fiscale e contributiva anno 2024).

[12] "*Gli approcci probabilistici all'analisi del rischio possono oggi avvalersi di un'ampia gamma di metodi e tecnologie che rendono possibile elaborare una grande mole di dati (big data) in tempi sempre più ridotti e, quindi, compatibili con le esigenze operative dell'Amministrazione fiscale*". (Cit. relazione sull'economia non osservata e sull'evasione fiscale e contributiva anno 2024).

fondimenti istruttori, risultando indicative di potenziali comportamenti fiscalmente illeciti.

2.1. Variabili di profilazione

Grazie alle analisi probabilistiche è possibile individuare le cc.dd. "variabili di profilazione" (variabili desunte o derivate) che sintetizzano le informazioni disponibili tramite valori segnaletici del livello di rischio fiscale che caratterizza i soggetti analizzati; tali variabili, così come precisato dall'articolo 22 del GDPR[13], non possono, però, essere utilizzate per determinare, in maniera totalmente automatizzata, pretese impositive per i soggetti individuati, salve eventuali eccezioni – attualmente non presenti nel nostro ordinamento –, che dovrebbero essere introdotte da norme ad hoc.[14]

Sul tema si richiama l'istituto del concordato preventivo biennale, introdotto nell'ordinamento dal d.lgs. n. 13 del 2024, in precedenza citato, per il quale è previsto che la relativa proposta "è

[13] "*L'interessato ha il diritto di non essere sottoposto a una decisione basata unicamente sul trattamento automatizzato, compresa la profilazione, che produca effetti giuridici che lo riguardano o che incida in modo analogo significativamente sulla sua persona*". Cfr. Comunicato stampa dell'Agenzia delle entrate del 17 dicembre 2025 che, richiamando una policy interna, chiarisce che "i sistemi di IA di tipo generativo non devono essere mai impiegati per la produzione diretta di atti amministrativi relativi all'attività istituzionale, anche di natura negoziale e della gestione delle controversie, sia in sede giudiziaria che extragiudiziale, nonché in fase endoprocedimentale".

[14] Nel documento di ricerca "*Intelligenza artificiale e accertamento tributario*" della Fondazione Nazionale Commercialisti del 18 novembre 2024 viene ritenuto che "*Presidio essenziale del diritto di difesa è l'obbligo di motivazione dell'atto intesa sia come individuazione delle ragioni giuridiche e dei presupposti di fatto ma anche come specificazione dei mezzi di prova (Art. 7 l. 212/2000): solo una motivazione intellegibile può dare la possibilità al contribuente di contestare efficacemente l'atto e quindi di orientare la decisione amministrativa o giurisdizionale verso la giusta imposta*".

elaborata dall'Agenzia delle entrate, in coerenza con i dati dichiarati dal contribuente e comunque nel rispetto della sua capacità contributiva, sulla base di una metodologia che valorizza, anche attraverso processi decisionali completamente automatizzati di cui all'articolo 22 del regolamento (UE) 2016/679, del Parlamento europeo e del Consiglio, del 27 aprile 2016, le informazioni già nella disponibilità dell'Amministrazione finanziaria, limitando l'introduzione di nuovi oneri dichiarativi"[15].

Si osserva altresì che anche i dati dei controlli fiscali effettuati dalle competenti articolazioni dell'Amministrazione finanziaria potrebbero partecipare significativamente alla fase di addestramento degli algoritmi di analisi predittiva, in modo tale che si possa, sulla base delle peculiari caratteristiche dei soggetti interessati dai controlli, elaborare punteggi degli indicatori utili all'identificazione e all'ordinamento dei profili di rischio; ad esempio potrebbe essere

[15] La medesima disposizione statuisce anche che "*La predetta metodologia, predisposta per i contribuenti di cui agli articoli 10, comma 1, e 23, comma 1, con riferimento a specifiche attività economiche tiene conto degli andamenti economici e dei mercati, delle redditività individuali e settoriali desumibili dagli indici sintetici di affidabilità fiscale di cui all'articolo 9-bis del decreto-legge 24 aprile 2017, n. 50, convertito, con modificazioni, dalla legge 21 giugno 2017, n. 96, e delle risultanze della loro applicazione, nonché degli specifici limiti imposti dalla normativa in materia di tutela dei dati personali. La metodologia è approvata con decreto del Ministro dell'economia e delle finanze, sentito il Garante per la protezione dei dati personali.*

2. Ai fini dell'elaborazione della predetta proposta, l'Agenzia delle entrate, oltre ai dati di cui al comma 1, ne acquisisce ulteriori dalle banche dati nella disponibilità dell'Amministrazione finanziaria e di altri soggetti pubblici, escluse quelle soggette alla disciplina di cui al decreto legislativo 18 maggio 2018, n. 51.

Con il decreto di cui al comma 1 sono individuate le specifiche cautele e le garanzie per i diritti e le libertà dei contribuenti di cui all'articolo 22, paragrafo 3, del regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio, del 27 aprile 2016, nonché le eventuali tipologie di dati esclusi dal trattamento. L'Agenzia delle entrate elabora e comunica la proposta attraverso i programmi informatici di cui all'articolo 8".

possibile associare le posizioni soggettive di interesse per il controllo e gli effetti attesi di un'eventuale attività accertativa, con l'obiettivo di incrementare l'efficienza e l'efficacia dell'operato degli uffici dell'Amministrazione finanziaria.

Con riferimento, invece, all'utilizzo delle informazioni "pubblicamente disponibili" (*i. e.*, quelle presenti sui social o sul web), si osserva che il Garante per la protezione dei dati personali le ha ritenute prive dei necessari requisiti di affidabilità e raccolte per finalità diverse da quelle sottese al trattamento in argomento (parere del Garante n. 3 dell'11 gennaio 2024); di conseguenza non partecipano nelle fasi di addestramento degli algoritmi nell'analisi del rischio relativa al contrasto agli illeciti fiscali.

Da ultimo si osserva che è possibile integrare le moderne tecniche di machine learning e di intelligenza artificiale con le metodologie di tipo statistico-econometrico, in modo da garantire sia la flessibilità in sede di modellazione dei dati sia l'interpretabilità dei risultati.

Finito di stampare nel mese di settembre 2026
presso Grafica Elettronica, Napoli

ISBN 979-12-235-0777-3

€ 24,00