

## **Intelligenza artificiale, scienze neuro-cognitive e diritto: metodologie e percorsi della ricerca**

Maria Novella Campagnoli

*Università degli Studi di Roma Tor Vergata*

Lo *special issue* dedicato a *Intelligenza artificiale e scienze neuro-cognitive nel diritto: dalla simbiosi alla sostituzione* – per scelta dei curatori – si articola in varie sezioni, nelle quali i contributi, pur essendo connotati tutti da una spiccata originalità tematica e da un approccio improntato alla multidisciplinarietà, si fondono coralmemente, armonizzandosi in maniera sinergica e complementare. Il risultato è un interessante e composito itinerario di ricerca, nel quale, le linee di indagine, di volta in volta proposte, coniugano felicemente la pluralità argomentativa con il rigore metodologico. Dalla ripresa delle questioni teoriche di ampio respiro si passa alla trattazione e alla problematizzazione di questioni più settoriali concernenti l'amministrazione della giustizia, il processo e la pena; si riserva, poi, ampio spazio alla disamina di temi specifici ed emergenti, come il rapporto uomo/macchina, la necessità di ripensare la soggettività, il mutamento della socialità e il benessere psicologico, per terminare con una riflessione sulla tutela dei (nuovi) diritti.

i) La prima sezione – *IA e pensiero giuridico: questioni epistemologiche* – si apre con il contributo di Lucio Franzese (*Sulla insufficienza dell'Intelligenza Artificiale e della legge*), in cui l'A., traendo le mosse da alcune considerazioni generali sull'uso dell'intelligenza artificiale, ne mette in luce le criticità nella sfera giuridica e sanitaria, come pure le tante difficoltà che si incontrano nell'elaborazione di una *governance* efficace.

Segue il saggio di Francesco Recanati (*L'Intelligenza Artificiale e il senso del diritto*), nel quale – a partire dal nesso costitutivo che congiunge la dimensione giuridica e quella antropologica – si valorizza il contributo che la cultura giuridica e il diritto possono offrire all'attuale società in cambiamento, affinché il rapporto fra l'uomo e i suoi strumenti non ne risulti del tutto stravolto.

Entro la prima macro-direttrice di analisi si colloca anche il saggio di Francesco Romeo (*La produzione culturale dell'artificiale: prodromi di uno studio metaepistemologico su diritto ed informazione*), secondo il quale, oggi, anche la cultura umana appare, a suo modo, del tutto diversa da un tempo. Questo è dovuto al fatto che la produzione della cultura stessa ha iniziato a vedere in prima linea

agenti e fattori non-umani. Di qui, l'urgenza di tornare a interrogarsi sulla verità e sul ruolo che l'informazione ha nella costruzione della realtà.

Chiude la sezione il lavoro di Paolo Savarese (*Diritto, sapere e intelligenza artificiale: una traccia epistemica*), in cui, riprendendo gli autorevoli studi sull'algebra di Francesco Maria Piccari, si problematizzano lo statuto della scienza giuridica, quello dello strumento normativo e, in generale, quello del sistema giuridico nel suo complesso.

*ii)* La seconda sezione – *IA generativa fra nuovi scenari e rischi etico-giuridici* – vede, per primo, il saggio di Agata C. Amato Mangiameli (*Intelligenza artificiale generativa: brevi note*). Qui l'A., dopo aver chiarito immediatamente la differenza fra IA e IA generativa, si sofferma sulla disamina dei vantaggi e dei rischi connessi a quest'ultima, riflettendo in chiave critica sulle possibili violazioni dei diritti e guardando, con particolare attenzione, soprattutto alla più recente regolamentazione introdotta a seguito dell'approvazione dell'*IA Act*.

Alle possibili implicazioni dell'IA si indirizza anche Stéphane Bauzon (*La place de l'Intelligence Artificielle (IA) dans le monde du droit*), a detta del quale – nonostante i tantissimi e strabilianti progressi – l'IA presenta ancora dei limiti e delle potenziali falle, in maniera particolare, in ambito giuridico, dove la complessità delle decisioni trascende spesso le capacità predittive degli algoritmi.

La sezione prosegue con il lavoro di Massimo Farina (*Degenerazione e rischio creativo dell'Intelligenza Artificiale "forte": forme di prevenzione e tutela complementare*), che mette in luce il (non remoto) pericolo di degenerazione dell'IA. Trattasi, come spiega l'A., del rovescio della medaglia della complessità tecnica degli attuali sistemi autonomi; una conseguenza con la quale, forse molto presto, saremo costretti a fare i conti.

In sostanziale continuità, si colloca il saggio di Francisco José Santamaría Ramos (*Implicaciones éticas y legales relacionadas con el uso creciente de la IA*), che sottolinea con forza l'importanza di un approccio responsabile ed etico all'IA. In modo particolare, guardando alla Dichiarazione di Montreal e alla Raccomandazione dell'UNESCO, Ramos riflette sulla regolamentazione e sul ruolo che l'etica dovrebbe avere nella progettazione e nello sviluppo dei sistemi di IA.

*iii)* La terza sezione – *Implicazioni decisionali e ricadute processuali* – che si orienta all'approfondimento del nesso fra IA e giustizia, processo, pena – annovera al suo interno anche vari lavori di diritto positivo. La trattazione ha inizio con il contributo di Fabrizio Corona (*L'evoluzione del processo decisionale tra giudizio automatico e sostituzione giudiziaria*), che si sofferma sulla formazione della decisione giudiziaria. Affiorano i dilemmi etici e giuridici associati al c.d. giudizio automatizzato, con un'attenzione particolare per il nesso fra intelletto umano e processo decisionale, nonché per la ricerca della perfezione nella decisione giudiziaria.

Nelle more di questa stessa macro-direttrice si iscrive anche il saggio di Chiara Grieco (*Tra simbiosi e condizionamento IA-centrico. Il divorzio fra legittimazione tecnica e agency umana*), che, richiamandosi a Friedman, riflette sulla simbiosi uomo/IA nell'ottica della transizione tecnologica del sistema giudiziario.

Si procede, poi, con il lavoro di Michele Martoni (*Spiegare le pieghe. Alcune osservazioni su trasparenza e spiegabilità*), nel quale si sottolinea la necessità di comprendere il funzionamento degli algoritmi e di scrutare l'opacità della *black-box*, come prerequisite imprescindibile rispetto alla spiegabilità e all'affidabilità stessa dell'IA.

L'indagine continua con Vincenzo Bruno Muscatiello (*Ex Machina. Il diritto penale, all'improvviso*), che prospetta un'interessante riflessione sul rapporto fra IA e istanze punitive, mettendo in evidenza taluni aspetti della realtà artificiale nei confronti dei quali il diritto penale "tradizionale" potrebbe risultare del tutto impreparato.

Nel medesimo solco tematico, anche Lorenzo Pulito (*Algoritmi predittivi e valutazioni prognostiche per la sostituzione della pena nella prospettiva della symbiotic artificial intelligence*) che – dinnanzi all'ambizione dell'IA di prevedere il futuro comportamento del reo, supportando, così, l'attività del giudice – smorza l'entusiasmo per il nuovo che avanza, mettendo in guardia nei confronti del possibile asservimento dell'essere umano al "regime algoritmico". L'idea che si prospetta (anche nel contesto delle nuove pene alternative e delle procedure recentemente introdotte dalla riforma Cartabia) è quella di un paradigma di interazione tra "intelligenze" (umana/artificiale) ispirato alle relazioni simbiotiche che si rintracciano in natura.

La sezione si completa con il saggio di Tomasz Snarski (*"Merciful Justice" and the Challenges of Posthumanism in Legal Philosophy: Some Comments on the Criterion for Evaluating Technological Progress in the Criminal Justice System*), a detta del quale, il criterio più autorevole per valutare lo sviluppo e l'uso dell'IA nell'ambito del diritto penale dovrebbe essere quello che si ispira alla c.d. "giustizia misericordiosa" (espressione di un diritto penale che è incentrato sull'uomo, sui suoi valori e sul suo ruolo all'interno del sistema giuridico stesso). È in questa prospettiva, a detta dell'A., si dovrebbe guardare alla sfida dell'integrazione dell'IA nel sistema della giustizia penale.

*iv*) La quarta sezione – *Neuroscienze, s/oggettività e modelli relazionali* – si apre con il saggio di Salvatore Amato (*Macchinazioni. Soggettività, agentività ed entità esotiche*), che, riprendendo il problema dell'opacità dell'IA, offre al lettore un'indagine sul confine tra la "meccanizzazione della mente" e "l'umanizzazione della macchina", giungendo a domandarsi se, forse, il richiamo – oggi sempre più frequente e diffuso – ad un'etica del *design*, ad un'info-etica, ad un'etica algoritmica, non faccia altro che reintrodurre (con variazioni e adattamenti) i principi e i valori della tradizione del diritto naturale.

Sempre in questa sezione, il contributo di Sveva D'Agosto (*La costruzione della figura femminile fra neuroni specchio e intelligenza artificiale simbiotica*) che, guardando alla costruzione del sistema simbiotico (intelligenza umana/artificiale), ragiona sull'implementazione delle rappresentazioni inautentiche, che sembrerebbero avere delle considerevoli ripercussioni sul funzionamento dei neuroni specchio. In questo contesto, ad emergere è una pericolosa deriva verso l'ingiustizia epistemica nella raffigurazione del femminile online, che concorre a generare dinamiche discriminatorie e un diffuso allontanamento dal riconoscimento e dall'affermazione dei diritti fondamentali delle donne.

Segue il lavoro di Teresa María García-Berrio (*Ética de la virtud para una IA compasiva*), che sviluppa delle considerazioni etiche circa l'impatto che l'IA ha sul tessuto sociale delle comunità contemporanee e sul benessere umano, richiamando l'attenzione sulla necessità di recuperare (e di riaffermare) valori come la gentilezza, l'empatia e la compassione.

Paola B. Helzel e Michele Leonetti (*Machine learning e depotenziamento dell'umano: quale futuro*) presentano una riflessione sul *Machine Learning* (ovvero su quel meccanismo di autoapprendimento in virtù del quale le macchine possono ricevere dati e utilizzarli per migliorarsi). Un ambito in continua evoluzione, che solleva parecchie riflessioni sulla ridefinizione (e sul confine) delle capacità macchiniche e umane.

Il percorso di analisi continua con l'articolo di Olimpia G. Loddo (*L'agire sociale ai tempi dell'intelligenza artificiale. Il concetto di "istituzione sociale-digitale"*). Qui, il concetto di istituzioni socio-digitali – inizialmente proposto da Teubner e Backers – viene visto come una possibile risorsa concettuale per affrontare le questioni poste dall'IA con riferimento alle dinamiche sociali. In particolare, lo studio guarda al ruolo dell'istituzione digitale nei sistemi sociali e a come la comunicazione mediata (o prodotta dall'IA) può concorrere alla sua ridefinizione.

Ad arricchire l'indagine, il saggio di Paolo Moro (*Persona elettronica. Una finzione giuridica per l'intelligenza artificiale*), nel quale – dopo aver proceduto all'esame della soggettività robotica (e dei parallelismi epistemologici tra la macchina calcolatrice e la coscienza umana) e dopo aver criticato la tendenza ad umanizzare le macchine (che verte su una concezione estremamente riduttiva della soggettività umana) – ci si interroga sulla possibilità di prevedere una nuova "personalità elettronica".

Con il contributo di Guido Saraceni (*La IA(D). Intelligenza Artificiale e benessere psicologico*) la ricerca si apre alla speculazione psicologica e, in modo particolare, alla disamina sulle conseguenze che i modelli di IA possono avere sul benessere psicologico dell'uomo. Lo scopo dello studio è quello di far luce sull'intricata relazione tra IA e salute mentale, mettendone in evidenza tanto i potenziali vantaggi, quanto i possibili limiti etici, sociali e tecnici.

La sezione termina con il lavoro di Maurizio Sozio (*Il deficit fonetico dello "Jus ex machina"*), in cui – preso atto dell'avvenuta realizzazione della nota

profezia di Leibniz (a detta del quale, un giorno, di fronte a una controversia, le parti avrebbero potuto sedersi e affidarsi alla risoluzione tramite un calcolo) – ci si chiede se un’entità calcolatrice, un’intelligenza artificiale, possa davvero sostituire l’uomo e il suo giudizio.

v) La quinta e ultima sezione – *IA e nuovi diritti* – comincia con il contributo di Wanda D’Avanzo (*Intelligenza artificiale e sanità digitale. Il diritto by design e il ruolo del legal design per la protezione dei dati personali*) che tocca un tema fra i più noti e significativi: l’interazione fra IA e sfera sanitaria e la necessaria tutela dei dati personali. In questo contesto, l’A. si sofferma sul consenso al trattamento, sul c.d. *legal design*, come pure sulle disposizioni più significative del GDPR.

L’attenzione nei confronti dei diritti si accresce grazie al lavoro di Laura Fabiano (*Algorithmic Recommendation ed Artificial Intelligence nell’esperienza statunitense ed europea fra esigenze strategiche e tutela dei diritti fondamentali*), che prospetta una comparazione fra gli approcci normativi al digitale che sono stati adottati negli Stati Uniti e in Europa – ma con uno sguardo anche al contesto cinese –, nei quali si rinvengono ancora parecchie difformità (i primi, Stati Uniti, seguitano a riconoscere la priorità alla crescita economica della nazione; la seconda, l’Europa, attribuisce da sempre la preminenza ai diritti fondamentali).

Decisamente interessante, è, poi, lo studio di Paolo Sommaggio e Shan Ali (*The Role of Blockchain in Managing Neuroscientific Evidence and Its Impact on Cognitive Liberty*), che, guardando alle neuroscienze e al neurodiritto, ipotizza la necessità di prevedere una sorta di “libertà cognitiva” (come nuova forma di tutela) da includersi in una “Dichiarazione dei Neurodiritti Umani”. Ciò, a partire dalla consapevolezza che la libertà cognitiva può rappresentare la nuova “arma concettuale” per difendere i diritti umani individuali dai neuro-paradigmi che, di fatto, li svuotano e li riducono a nulla.

In conclusione, il saggio di Radosław Strugała (*Liability for Damage Caused to the Data Subject in the Event of GDPR Violation “by Artificial Intelligence”: Between Legal Discrepancies and Policy Questions*) affronta l’importante questione della responsabilità per violazione dei dati da parte dell’IA, ipotizzando, fra l’altro, una possibile interazione fra l’applicazione dell’articolo 82 del GDPR e un regime speciale di responsabilità pensato per far fronte ai danni causati dall’IA.

Assieme a Piero Marra, desidero ringraziare, in prima battuta, i Direttori della Rivista *L’Ircocervo* (Agata C. Amato Mangiameli, Paolo Becchi, Gian Pietro Calabrò, Lucio Franzese, Antonio Incampo, Torquato G. Tasso), che, con grande generosità e fiducia, ci hanno concesso l’onore di curare l’edizione di questo *special issue*. Un ringraziamento particolare va, ovviamente, anche ai numerosi autori che hanno deciso di proporre un contributo, nonché a tutti i revisori anonimi (italiani e stranieri), il cui operato, particolarmente attento e scrupoloso, è stato essenziale per garantire una selezione imparziale dei saggi e per assicurare l’elevata qualità scientifica del fascicolo. Infine, non posso non esprimere la mia gratitudine verso Piero, collega e compagno in questa bellissima avventura editoriale.