



Collana del Dipartimento di Giurisprudenza
dell'Università di Milano-Bicocca

Silvia Salardi

Intelligenza artificiale e semantica del cambiamento: una lettura critica



G. Giappichelli Editore – Torino



Collana del Dipartimento di Giurisprudenza
dell'Università di Milano-Bicocca

Direttore:

Loredana Garlati

Comitato editoriale:

Maurizio Arcari

Chiara Buzzacchi

Paolo Casella

Adolfo Ceretti

Diana Cerini

Giovanni Chiodi

Filippo Danovi

Monica Delsignore

Giovanni Guglielmetti

Costanza Honorati

Giovanni Iorio

Gioacchino La Rocca

Natascia Marchei

Alfredo Marra

Claudio Martinelli

Oliviero Mazza

Stefania Ninatti

Claudia Pecorella

Andrea Rossetti

Carlo Ruga Riva

Nicola Sartori

Franco Scarpelli

Antonello Tancredi

Giulio Vigevani

Silvia Salardi

Intelligenza artificiale e semantica del cambiamento: una lettura critica



G. Giappichelli Editore – Torino

© Copyright 2023 - G. GIAPPICHELLI EDITORE - TORINO

VIA PO, 21 - TEL. 011-81.53.111 - FAX 011-81.25.100

<http://www.giappichelli.it>

ISBN/EAN 978-88-921-2465-3

ISBN/EAN 978-88-921-7779-6 (ebook - pdf)

Questo volume è pubblicato con il contributo del Dipartimento di Giurisprudenza dell'Università degli Studi di Milano-Bicocca.

L'opera ha ottenuto la valutazione positiva di due revisori anonimi, secondo il procedimento previsto dal Regolamento della Collana, consultabile sul sito del Dipartimento di Giurisprudenza.

Stampa: Stampatre s.r.l. - Torino

Le fotocopie per uso personale del lettore possono essere effettuate nei limiti del 15% di ciascun volume/fascicolo di periodico dietro pagamento alla SIAE del compenso previsto dall'art. 68, commi 4 e 5, della legge 22 aprile 1941, n. 633.

Le fotocopie effettuate per finalità di carattere professionale, economico o commerciale o comunque per uso diverso da quello personale possono essere effettuate a seguito di specifica autorizzazione rilasciata da CLEARedi, Centro Licenze e Autorizzazioni per le Riproduzioni Editoriali, Corso di Porta Romana 108, 20122 Milano, e-mail autorizzazioni@clearedi.org e sito web www.clearedi.org.

A mia madre

INDICE

	<i>pag.</i>
<i>Ringraziamenti</i>	IX
<i>Premessa</i>	XI

CAPITOLO I

PROGRESSO TECNOLOGICO E QUESTIONE SEMANTICA: *DIFFINITIO E PARTECIPATIO*

1. Introduzione	1
2. L'avvento dell'Intelligenza Artificiale e la semantica del cambiamento	5
3. <i>Diffinitio quid nominis</i> v. <i>diffinitio quid rei</i>	18
4. Tassonomia concettuale dell'antropomorfizzazione dell'IA	26
4.1. Intelligenza, autonomia e imprevedibilità della persona umana	28
4.2. Intelligenza, autonomia e imprevedibilità della persona artificiale	35
5. La questione semantica come problema etico	39
6. La questione semantica come questione costituzionale e democratica	47
7. <i>Trustworthy AI</i> : il controllo linguistico come presupposto per costruire un rapporto di fiducia con le tecnologie avanzate	50

CAPITOLO II

PROGRESSO TECNOLOGICO E QUESTIONE TECNICA: LA SEMANTICA DEL *POSSE*

1. Introduzione	55
2. Breve cronistoria dell'IA	56
3. Perché i problemi tecnici sono problemi etici?	58
4. La conoscenza nella prospettiva tecnologica e in quella scientifica	67
5. La questione tecnica e la semantica dell'antropomorfizzazione: dominio e potere sull'individuo	77
6. Il problema tecnico che dissimula un problema socioculturale	78

CAPITOLO III

PROGRESSO SCIENTIFICO E TECNOLOGICO
E QUESTIONE GIURIDICA: LA SEMANTICA DEL *LICERE*

1. Introduzione	83
2. Che rapporto è possibile tra diritto, scienza e tecnologia?	84
2.1. Tecnologia come oggetto di regolazione	87
2.2. Tecnologia come agente regolatore	89
3. Categorie giuridiche alla prova dell'IA	94
4. I diritti fondamentali sono la cornice adeguata a regolare l'innovazione tecnologica? Il caso del riconoscimento facciale	101
4.1. Il principio di precauzione e la tutela effettiva dei diritti	106
4.2. Il principio di esplicabilità e i diritti portati a compimento	107

CAPITOLO IV

CASO DI STUDIO: IA E *BIG DATA* IN MEDICINA

1. Introduzione	113
2. I dati e loro impiego nella pratica clinica e nella ricerca scientifica	115
3. IA e relazione di cura: terzo <i>super partes</i> ?	120
3.1. La relazione di cura tra diritti e doveri	121
3.2. La relazione di cura mediata dall'IA	124
4. Neurotecnologie e potenziamento umano: le nuove frontiere della medicina	128
4.1. Il biopotenziamento in breve	130
4.2. Profili etico-giuridici del potenziamento neurotecnologico	133
4.3. Neurotecnologie e ideologia antropomorfizzante: alcune riflessioni conclusive	135

Bibliografia

Ringrazio la Professoressa Patrizia Borsellino per le osservazioni sulla prima versione di questo volume.

PREMESSA

La letteratura in tema di Intelligenza Artificiale (IA), di digitalizzazione della società e, più in generale, di tecnologie emergenti è vasta. Ci si può, pertanto, chiedere per quale motivo proporre un altro volume su queste tematiche. I motivi possono essere di diversa natura e rilevanza. Richiamando quello più scontato, possiamo dire che il filosofo del diritto non può sottrarsi dal proporre riflessioni su un fenomeno che non solo è sulla bocca di tutti, ma che è destinato a pervadere e modificare sia l'organizzazione sociale sia le relazioni umane. Proprio la radicalità dei cambiamenti sottesi alla società digitalizzata richiede al filosofo del diritto, in particolare di indirizzo analitico-linguistico, di approfondire la questione semantica che si accompagna a tale trasformazione. La narrazione etico-giuridica dell'IA si svolge, infatti, grazie all'impiego di concetti volti a orientare la visione dei cambiamenti sociali e relazionali, che caratterizzeranno il prossimo futuro. Con particolare riferimento al contesto europeo, la narrazione predilige termini ed espressioni linguistiche volte a consolidare la "visione incentrata sulle persone" (*human-centric vision*), richiamata in tutti i documenti normativi che, da qualche anno, contribuiscono a costruire il percorso europeo su questo tema.

La questione semantica non è quindi intesa, in questa sede, come pura disquisizione sulla preferibilità di un significato rispetto ad un altro per mera preferenza estetica. Al contrario, in questo volume si affronterà la questione semantica in vista di un duplice scopo: per un verso, «attraverso il chiarimento dei significati e della logica del linguaggio, aiutare la gente a pensare»¹, per altro e connesso verso, mettere in luce il lavoro "costruttivo" che passa attraverso il controllo linguistico, volto a restituire funzionalità ed efficienza a uno dei principali strumenti di relazione di cui gli esseri umani dispongono. Entrambi gli obiettivi possono giocare un ruolo chiave nell'orientare il percorso etico della trasformazione digitale e dell'impiego dell'IA in moltissimi, per non dire quasi tutti, i settori in cui gli esseri umani vivono e sviluppano la loro personalità.

Il chiarimento linguistico è pertanto volto a evidenziare e superare opacità, equivoci e contraddizioni, sia intenzionali sia involontarie, che possono sorgere

¹ U. SCARPELLI, *Filosofia analitica, norme e valori*, Milano, Edizioni Comunità, 1962, p. 12.

con l'uso di concetti in determinate argomentazioni pratiche, portate a sostegno di specifiche linee d'azione che influenzano il processo decisionale sul piano pubblico. Linee di azione che rischiano di spingere lo sviluppo tecnologico verso un'unica destinazione e una ben precisa visione dell'ordine sociale a cui i soli valori sottesi sono efficienza e profitto. L'analisi della questione semantica non è l'unica necessaria per trovare argini a questa eventuale deriva. Essa rappresenta però un passo imprescindibile perché porta con sé la possibilità di sgomberare il campo da grossolani equivoci, che sono spesso alla base dell'apparente impossibilità di trovare un comune denominatore su cui costruire buone soluzioni pratiche per gli individui nel momento storico in cui vivono e, in prospettiva, per le generazioni future.

La riflessione sul linguaggio usato nel discorso normativo, che accompagna l'innovazione in materia di IA, è qui intesa come un utile strumento per trovare risposte o quantomeno indirizzi condivisi in relazione alle problematiche etiche che tale sviluppo solleva. L'analisi della questione semantica può, invero, ben orientare la ricerca di ragioni giustificanti per eventuali strade alternative a quelle proposte da chi detiene il potere sulla tecnologia e altresì economico, ossia soluzioni che tutelino effettivamente la dimensione egualitaria della distribuzione dei benefici di tale progresso. In sintesi, attraverso l'analisi linguistica ci si propone di tenere alta l'attenzione sugli usi ideologici delle scoperte scientifiche, delle conversioni tecniche di tali scoperte e più in generale del discorso sulla tecnologia.

A fronte di questa premessa metodologica, è opportuno accennare in questa sede le tre direttrici principali lungo le quali si articolerà l'analisi della questione semantica in questo volume. Va innanzitutto anticipato che la riflessione proposta non ha né la pretesa di essere esaustiva né risolutiva. Il più modesto scopo è quello di portare alla luce ciò che, rimanendo tra le pieghe linguistiche, può essere utilizzato in modo ubiquo per servire opposti scopi e raggiungere fini e affermare valori non necessariamente eticamente sostenibili. In altre parole, sollevare la questione semantica sarà utile per rinforzare la visione sistemica della «logica che permea la tecnologia e la trasforma in azione»², che autori come Shoshana Zuboff ed Éric Sadin, per citarne solo alcuni³, propongono in merito a molti accadimenti relativi agli indirizzi seguiti dallo sviluppo tecnologico tra la metà degli anni '90 del secolo scorso e gli inizi del nuovo secolo. Accadimenti

² S. ZUBOFF, *Il capitalismo della sorveglianza. Il futuro dell'umanità nell'era dei nuovi poteri*, Roma, LUISS University Press, 2019, p. 25.

³ Si veda anche A. PUNZI, *Difettività e giustizia aumentata. L'esperienza giuridica e la sfida dell'umanesimo digitale*, *Ars Interpretandi*, 2021, 1, pp. 113-128. Cfr. anche F. PASQUALE, *Le nuove leggi della robotica. Difendere la competenza umana nell'era dell'intelligenza artificiale*, Roma, LUISS University Press, 2021; H. NOWOTNY, *Le macchine di Dio. Gli algoritmi predittivi e l'illusione del controllo*, Roma, LUISS University Press, 2022.

che, sebbene in origine non fossero necessariamente collegati fra loro e nemmeno frutto di una programmazione intenzionale – anzi sembravano orientati verso “la difesa del pubblico”⁴ – sono venuti a costituire un fenomeno pervasivo e con una sua progettualità, che Sadin denomina *tecnologia integrale*⁵ con una *potenza ingiuntiva*⁶. Per poter procedere in questa direzione le direttrici poc’anzi richiamate, ma non ancora esplicitate, saranno le seguenti.

La prima linea di riflessione, che sarà affrontata nel primo capitolo del volume, si soffermerà sulla questione semantica nella sua dimensione definitoria (*diffinitio*), seguendo la tradizione della filosofia analitico-linguistica⁷. Assumendo come rilevante il ruolo dello strumento definitorio nella chiarificazione dei principali termini ricorrenti nel dibattito europeo attorno all’IA in merito alle scelte normative in questa materia, si cercherà di mettere in luce, attraverso l’individuazione delle relazioni tra le diverse nozioni che si prenderanno in considerazione, quanto e come la questione semantica nella sua accezione definitoria sia propedeutica ad affrontare due questioni: 1) la realizzazione dei c.d. valori europei; 2) la comunicazione delle informazioni relative alle trasformazioni in atto, funzionale all’effettiva ed attiva partecipazione di tutti i soggetti interessati dalle trasformazioni tecnologiche, indipendentemente dai loro specifici ruoli (produttore, consumatore, programmatore, finanziatore ecc.) per il conseguimento della *visione incentrata sulle persone*.

La seconda direttrice, lungo la quale proseguirà il percorso di analisi nel secondo capitolo, indaga la questione semantica derivante dagli aspetti tecnici che delimitano attualmente l’ambito della fattibilità (*posse*) e operatività di certe tecnologie: dal problema degli algoritmi *black-box*, al tema dell’accesso effettivo ai *dataset*, fino alla questione dei possibili *biases* presenti nei dati. A molte delle questioni tecniche, che verranno richiamate, sono infatti strettamente correlate scelte operative, anche sul piano giuridico, che chiamano in causa nozioni come trasparenza, intellegibilità ed altre, il cui significato è tutt’altro che unanimemente condiviso. La dimensione del *posse* incontra quello delimitato dalle questioni di *licere*, anche e soprattutto, attraverso le scelte concettuali e terminologiche che possono alternativamente accentuare o smorzare i toni circa i limiti tecnici di una data tecnologia, laddove operare nell’una o nell’altra direzione sia funzionale, ad esempio, a evitare che l’innovazione tecnologica incontri eccessivi aprioristici ostacoli nel suo potenziale sviluppo.

La terza linea di analisi chiamerà direttamente in causa gli aspetti giuridici

⁴ *Ibidem*, p. 55.

⁵ E. SADIN, *Critica della ragione artificiale. Una difesa dell’umanità*, Roma, LUISS University Press, 2019, p. 12.

⁶ *Ibidem*.

⁷ U. SCARPELLI, *La definizione nel diritto*, in ID. (a cura di), *Diritto e Analisi del Linguaggio*, Milano, Edizioni Comunità, 1976, pp. 183-197.

(*licere*) della trasformazione che la società sta vivendo e sarà oggetto del terzo capitolo. Seguendo la tradizione analitico-linguistica si può porre l'accento sull'analisi del diritto come linguaggio. Ciò consente di mettere in luce il suo stretto legame con il linguaggio ordinario e le trappole linguistiche indotte da questo legame. Per superare tali trappole si richiamerà il ruolo virtuosamente *performativo* che un'attenta selezione dei termini normativi può svolgere per evitare di indurre i consociati in fraintendimenti e confusioni. In aggiunta a questa funzione, il diritto e il suo linguaggio saranno considerati come una *costruzione interamente artificiale*⁸ le cui caratteristiche entrano in gioco non solo quando si tratta di trovare soluzioni adeguate in termini di tecnica normativa in funzione dei problemi interpretativi, ma altresì quando si tratta di orientare scelte di valore. La prospettiva del diritto come *costruzione artificiale* lascia al potere normativo la scelta se costruirlo come «legge del più forte»⁹ o come «legge del più debole»¹⁰, che pone limiti, vincoli e controlli nei confronti dei poteri privati e pubblici. Quest'ultimo aspetto è particolarmente rilevante nell'individuazione del limite oltre il quale «la logica economica diventa incompatibile con il rispetto della persona costituzionalizzata»¹¹. Nell'attuale scenario, c'è chi tende a proporre nuove strategie di regolazione espresse da suggestive formule come *Code is law*¹², ossia l'altra faccia della medaglia della formula *Law is code*. Cosa significano questi rovesciamenti dal punto di vista semantico nell'economia del discorso giuridico? Quali ricadute possono avere sull'equilibrio che nel contesto europeo si cerca di mantenere tra diversi interessi e diritti in gioco? Se, in Europa, i diritti fondamentali possono essere considerati non solo il fondamento formale del sistema giuridico, ma altresì l'elemento identitario, quali sono i meccanismi e gli strumenti più adeguati a mantenere l'innovazione tecnologica sul percorso dei diritti fondamentali? Il richiamo ai diritti fondamentali come stella polare delle decisioni circa le linee di azione non deve restare solo sulla carta. Per contrastare i pericoli profondi del nuovo paradigma antropologico che la trasformazione digitale, realizzata nell'ambito di un'economia globalizzata, può provocare, in particolare «il regresso cognitivo, il disprezzo della ragione e della scienza»¹³, i diritti fondamentali devono restare «una narrazione capace di unificare, di produrre relazioni, di rivelare la radice comune di iniziative che si ma-

⁸ L. FERRAJOLI, *Dei diritti e delle garanzie. Conversazione con Mauro Barberis*, Bologna, Il Mulino, 2013, p. 45.

⁹ *Ibidem*.

¹⁰ *Ibidem*.

¹¹ S. RODOTÀ, *Il diritto di avere diritti*, Roma-Napoli, Laterza, 2012, p. 414.

¹² L. LESSIG, *Code is Law. On Liberty in Cyberspace*, in *Harvard Magazine*, January-February 2000, <https://www.harvardmagazine.com/2000/01/code-is-law.html>.

¹³ Cfr. A. PUNZI, *Difettività e giustizia aumentata. L'esperienza giuridica e la sfida dell'umanesimo digitale*, cit. p. 117.

nifestano nei più diversi luoghi del mondo»¹⁴. Tuttavia, per valutare se il richiamo ai diritti fondamentali possa davvero rappresentare uno spartiacque tra una società egualitaria e una in cui il potere tecnologico ed economico è, nella prassi, *legibus solutus*, occorre ricordare che questi diritti, lungi dall'essere assoluti e fondati su una presunta immodificabile natura umana, «sono strumenti»¹⁵ che, per essere efficaci, hanno bisogno del rispetto di requisiti indispensabili¹⁶, ossia dell'affermazione, dell'attuazione e della giustiziabilità, ma altresì di un costante monitoraggio circa l'accettazione (introiezione) del loro rispetto come fine primario da raggiungere e, al contempo base giustificativa da cui partire per ispirare la costruzione di un nuovo ordine sociale da parte di tutti gli attori operanti nella società digitalizzata.

In sintesi, nel volume il lettore potrà seguire i cambiamenti, in parte già attuati e in parte da attuare, attraverso una lettura critica dell'uso dei concetti che accompagnano e modellano la narrazione della *Società dell'informazione*.

Silvia Salardi

¹⁴ S. RODOTÀ, *Il diritto di avere diritti*, cit. p. 416.

¹⁵ P. BORSELLINO, *Bioetica tra 'moralì' e diritto*, Milano, Raffaello Cortina, 2018, p. 105.

¹⁶ A. FACCHI, *Breve storia dei diritti umani*, Bologna, Il Mulino, 2007, p. 13.

CAPITOLO I

PROGRESSO TECNOLOGICO E QUESTIONE SEMANTICA: *DIFFINITIO* E *PARTECIPATIO*

SOMMARIO: 1. Introduzione. – 2. L'avvento dell'Intelligenza Artificiale e la semantica del cambiamento. – 3. *Diffinitio quid nominis* v. *diffinitio quid rei*. – 4. Tassonomia concettuale dell'antropomorfizzazione dell'IA. – 4.1. Intelligenza, autonomia e imprevedibilità della persona umana. – 4.2. Intelligenza, autonomia e imprevedibilità della persona artificiale. – 5. La questione semantica come problema etico. – 6. La questione semantica come questione costituzionale e democratica. – 7. *Trustworthy AI*: il controllo linguistico come presupposto per costruire un rapporto di fiducia con le tecnologie avanzate.

1. *Introduzione*

Il tema dell'Intelligenza Artificiale (IA) è oggi, unitamente alla questione ambientale, l'ambito più dibattuto nei diversi campi del sapere e quello a cui le istituzioni nazionali, europee e internazionali rivolgono la loro massima attenzione. In gioco, infatti, vi è la *sovranità* sulle tecnologie¹. Chi arriva primo nel controllo delle strutture informatiche e nella regolamentazione dell'uso delle tecnologie intelligenti potrà di fatto avvantaggiarsi molto sul piano economico, sul piano del potere di controllo delle attività umane a livello globale e sul piano della costruzione di un determinato modello e di una specifica visione della società. In gara ci sono le grandi potenze mondiali.

Nel contesto europeo, questa gara è stata presa molto sul serio, soprattutto per quello che riguarda la previsione di regole da apprestare, per un verso, per impedire che sul mercato europeo si possano immettere prodotti e sistemi di intelligenza artificiale in maniera selvaggia, con il correlato rischio di violazione dei diritti e delle libertà fondamentali e, per altro verso, per obbligare indiret-

¹Ursula von der Leyen, Presidente dell'Unione europea, ha sottolineato che l'Europa deve raggiungere la *sovranità tecnologica* (*technological sovereignty*) in settori chiave, cfr. EUROPEAN PARLIAMENT, *Digital Sovereignty for Europe*, EPBR Ideas Papers, 2020, accessibile a questo link [https://www.europarl.europa.eu/RegData/etdes/BRIE/2020/651992/EPRS_BRI\(2020\)651992_EN.pdf](https://www.europarl.europa.eu/RegData/etdes/BRIE/2020/651992/EPRS_BRI(2020)651992_EN.pdf).

tamente gli altri sfidanti in questa corsa a prevedere normative in linea con quelle europee (*extraterritorial reach*).

È con questo duplice obiettivo che l'Unione europea ha elaborato una Proposta di Regolamento sull'IA². Trattasi di un importante traguardo di cui avremo modo di occuparci nel prosieguo del lavoro.

Va qui notato che questo traguardo normativo è raggiunto in un contesto pubblico in cui la dialettica che definisce il rapporto tra i due mondi – umano e macchinico – è ancora alimentata da minacciose contrapposizioni sia nella rappresentazione mediatica al grande pubblico sia nella sfera pubblico-istituzionale. Non bisogna, infatti, andare molto indietro nel tempo per trovarne traccia giuridica. Basta richiamare alla mente il processo di elaborazione di una regolamentazione europea in materia di robotica, che è rimasto palesemente imbrigliato in questa dialettica, quando nel 2017 il Parlamento europeo propose il riconoscimento di personalità giuridica ai robot (*e-personality*)³.

In questo processo dialettico, che proponiamo di denominare *dialettica degli opposti*, si contrappongono, per un verso, l'uomo, il cui destino sembra tracciato verso una marginalità ineluttabile e, per altro verso, le macchine, destinate a sostituire l'essere umano in quanto – si assume – capaci di prestazioni umanamente irraggiungibili in pressoché tutte o quasi le attività. Alla base di questa dialettica vi è sottesa un'ideologia antropomorfizzante del macchinico. Con questa espressione si indica quel processo interpretativo *per relationem* tra l'essere umano e altri enti, che ha radici molto risalenti⁴. Tale processo interpretativo legge ogni tipo o specie di realtà *per relationem* con i comportamenti umani o per somiglianza o analogia con essi. Ciò vale sia per le storiche interpretazioni di Dio e/o della Natura in termini antropomorfi sia oggi per l'IA.

Questa modalità esplicativa soffre sia di un errore metodologico sia di uno epistemologico. Da un punto di vista metodologico, l'errore deriva dal fatto che la maggior parte degli studi per la costruzione di macchine intelligenti ha come riferimento il cervello umano sia con riguardo al suo funzionamento razionale sia emotivo. A fronte di questo stato di cose, appare naturale ragionare *per relationem*, anche quando il dibattito si sposta dal piano descrittivo a quello norma-

² Proposta di Regolamento del Parlamento europeo e del Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale e modifica di alcuni atti legislativi dell'Unione del 21 aprile 2021.

³ Risoluzione recante raccomandazioni alla Commissione europea concernenti norme di diritto civile sulla robotica, accessibile a questo link <https://eur-lex.europa.eu/legal-content/IT/ALL/?uri=Celex%3A52017IP0051>, ultimo accesso giugno 2022.

⁴ La visione antropomorfa del macchinico ha origine nel XII secolo in Europa, con una creatura mistica denominata golem, che aveva una forma simil-uomo, ma possedeva maggiore forza. Cfr. A. WÖLL, *Der Golem: Kommt der erste künstliche Mensch und Roboter aus Prag?*, in M. NEKULA, J. ROGALL (hrsg.), *Deutsche und Tschechen: Geschichte-Kultur-Politik*, München, Beck, 2001, pp. 235-245.

tivo. Ed è qui che si insinua il problema metodologico. Va, infatti, evidenziato che utilizzare le conoscenze sul funzionamento della biologia e della neurologia umana per poter costruire macchine in grado di svolgere compiti complessi, considerando l'essere umano come il modello biologico più avanzato a cui ispirarsi, non implica necessariamente derivare la conclusione normativa che tali macchine siano da equiparare agli umani anche sul piano valoriale (a meno di non compiere una fallacia naturalistica). Questo errore metodologico, che ha inficiato buona parte della storia della scienza, finché non si è liberata dalle visioni antropomorfe della natura ed ha potuto studiarla *juxta propria principia*⁵, vizia preconcettualmente anche il contemporaneo sviluppo tecnologico.

Sul piano epistemologico, con la spiegazione *per relationem*, la tendenza è quella di mettere in evidenza le differenze tra macchine intelligenti ed esseri umani, portando principalmente a esaltare i difetti dei secondi e le virtù delle prime. L'aspetto conoscitivo delle reali potenzialità e limiti delle macchine così come dell'uomo rimane così offuscato dalla spiegazione *per relationem* consentendo all'ideologia riduzionista, alimentata dal processo di antropomorfizzazione del macchinico, di rifiorire.

A fronte della permanenza e persistenza di questa dialettica tra opposti, che ha pervaso non solo la letteratura e la cinematografia a partire dal secolo scorso, ma che si è altresì radicata nel sentire comune e nelle istituzioni politiche europee, appare opportuno individuare gli elementi centrali di questa contrapposizione per tentare, laddove possibile, una conciliazione e/o un superamento di questo approccio dialettico.

Una delle principali fonti, sebbene non l'unica, a cui si abbevera questa dialettica, è il linguaggio. In particolare, fanno il gioco di questa dialettica l'ubiquità, la vaghezza e l'ambiguità di nozioni e termini che vengono impiegati negli argomenti delle tesi in lotta. È quindi sul linguaggio che ci soffermeremo per tentare di operare una ripulitura da elementi che appartengono all'ideologia antropomorfizzante e riduzionista, che ha accompagnato lo sviluppo dell'IA dal secolo scorso ad oggi. Tale ideologia è uscita dalla letteratura e dalla cinematografia e si è diffusa, tramite il linguaggio, nel sentire comune e, da lì, influenza anche diversi settori del sapere⁶, nonché i contesti istituzionali.

⁵ B. TELESIO, *De rerum natura juxta propria principia*, 1565, riproposta (a cura di), V. STAMPANATO, *Filosofi italiani*, collezione promossa dalla Società Filosofica Italiana, Genova, A.F. Formiggini, vol. II, 1913.

⁶ Ciò accade per vari motivi, ma nel caso, ad esempio, del linguaggio giuridico, la ragione principale per cui una parola può essere ivi usata secondo le regole che la definiscono nel linguaggio comune dipende dal prestito semiotico tra i due linguaggi. Molti dei problemi del linguaggio comune dipendono dall'«ingenua tendenza a ontologizzare le entità concettuali» da parte degli uomini, ma «il linguaggio comune [...] e quello giuridico, più son distinti, e meglio è. Quante volte si usano [...] i termini giuridici secondo le regole che li definiscono nel linguaggio comune, e non si sta in guardia contro la "carica" emotiva» di parole che, invece, richiedono una valutazione stret-

L'intento dei paragrafi che seguono è quello di mettere in luce come, attraverso concetti ambigui e vaghi, che possono costituire vere e proprie trappole linguistiche, si possa mantenere viva la dialettica degli opposti, che rappresenta una modalità surrettizia di esercizio del potere. Con tale approccio dialettico, infatti, si rischia di inficiare scelte razionali sul piano pubblico, ossia scelte che tengano conto dei reali vantaggi e svantaggi dell'innovazione mediante IA e dell'equa distribuzione dei benefici che essa può portare.

Consapevoli della duttilità del linguaggio, in particolare di quello ordinario, e della stratificazione di significati che si realizza durante le vicende storiche dei termini, quello che si propone è una riflessione sugli usi attuali di nozioni centrali nella narrazione pubblica e istituzionale attorno all'IA e su come questi usi possano foggare la costruzione delle politiche e delle visioni sul macchinico e sull'umano.

Per dirla con Uberto Scarpelli, il nostro scopo è di operare una «restituzione del linguaggio quale buon mezzo di espressione, comunicazione ed orientamento, buono non secondo i criteri interni al linguaggio, buono per l'uomo con i suoi attuali atteggiamenti, bisogni e progetti nelle circostanze in cui ora si trova»⁷. La scelta di soffermarsi sul linguaggio è determinata da due ragioni. La prima è che il linguaggio rappresenta il principale mezzo di comunicazione tra gli esseri umani e di influenza sul loro comportamento e la seconda è che le stesse tecnologie di cui discorreremo sono costruite sulla base di linguaggi. Un contesto in cui emerge, ad esempio, la centralità della questione linguistico-semantica è quella delle *Human Language Technologies*, anche denominate *language technologies* (LTs)⁸. L'impiego di queste tecnologie del linguaggio ha sollevato una serie di problemi etico-giuridici, quali la sottorappresentazione linguistica, il rischio di estinzione digitale dei linguaggi *underresourced* e, in generale, le barriere linguistiche che rischiano di non venire abbattute, ma rinsaldate. Le problematiche correlate alle LTs nel contesto europeo mostrano chiaramente che l'uso delle nuove tecnologie richiede un'attenta riflessione sul linguaggio. La riflessione può essere incentrata, come nel caso di specie delle LTs, su un'attenta valutazione dell'impiego del linguaggio mediante queste tecnologie per evitare discriminazioni e omologazioni discriminanti, che contrasterebbero con l'articolo 22 della Carta dei diritti fondamentali dell'Unione europea⁹.

tamente giuridica, cfr. U. SCARPELLI, *Scienza del Diritto e Analisi del Linguaggio*, in U. SCARPELLI, PAOLO DI LUCIA, M. JORI (a cura di), *Il Linguaggio del Diritto*, Milano, LED Edizioni Universitarie, 1994, pp. 87-93, ivi, p. 93.

⁷ U. SCARPELLI, *Filosofia analitica norme e valori*, Milano, Edizioni Comunità, 1962, p. 22.

⁸ PANEL FOR THE FUTURE OF SCIENCE AND TECHNOLOGY (STOA), *Language equality in the digital age. Toward a Human Language Project*, Brussels, 2017 accessibile al link [https://www.europarl.europa.eu/thinktank/en/document.html?reference=EPRS_STU\(2017\)598621](https://www.europarl.europa.eu/thinktank/en/document.html?reference=EPRS_STU(2017)598621), ultimo accesso luglio 2021. Lo studio chiarisce che le LTs «are software systems designed to handle human language in all its forms», p. 19.

⁹ L'articolo 22 intitolato *Diversità culturale, religiosa e linguistica* recita «L'Unione rispetta la

Ma l'ambito di analisi del linguaggio può estendersi, come cercheremo di fare in questo volume, ai concetti usati nella narrazione pubblica e istituzionale attorno all'IA. Questa proposta di approfondimento è volta a indagare due ordini di vicende intimamente connesse con le vicende del linguaggio e della sua forza sociale¹⁰, ossia «le trasformazioni dei soggetti umani e le trasformazioni delle circostanze intorno ai soggetti umani»¹¹. Sebbene il linguaggio non sia l'unico elemento a plasmare la rappresentazione della realtà, cionondimeno «le circostanze intorno agli uomini sono per gli uomini quello che sono perché gli uomini ne parlano e ci operano su impiegando un certo linguaggio e sarebbero diverse ove ne parlassero e ci operassero su con un linguaggio diverso»¹². Ciò a dire che decenni di dialettica degli opposti hanno profondamente influenzato non solo la visione che le persone hanno del macchinico, ma soprattutto l'idea, troppo spesso improntata al riduzionismo, che si sono fatti di sé stessi e della specie umana.

L'analisi dei concetti usati nel racconto al pubblico sull'IA sarà utile a riportare un equilibrio narrativo che faccia emergere visioni sociali e culturali alternative alle proposte fagocitanti di chi detiene il potere tecnologico ed economico.

2. *L'avvento dell'Intelligenza Artificiale e la semantica del cambiamento*

Molto si è scritto e si scrive sulla capacità dell'IA di rivoluzionare il nostro modo di vivere. Klaus Schwab parla di *quarta rivoluzione*¹³, intendendo con ciò una trasformazione caratterizzata da tecnologie intelligenti in grado di combinare la sfera fisica, quella digitale e quella biologica e che, per le loro caratteristiche, metteranno in discussione lo stesso significato di umano e di natura umana. Si assume che, per mezzo dell'IA, sarà possibile operare un mutamento radicale del nostro vivere e della nostra stessa natura. Questa assunzione, che sostiene tutta la narrazione dell'IA, ma che in quanto assunzione non implica necessariamente la verità della premessa che si assume, ossia che vi sarà un benessere egualmente distribuito in tutto il globo con l'intervento dell'IA in pressoché tutte le attività umane, viene divulgata attraverso precise scelte semantiche. Tra queste scelte spiccano due termini, quello di rivoluzione e quello di innovazione, che stanno alla base della semantica del cambiamento, che permea il nostro

diversità culturale, religiosa e linguistica». Il rispetto della diversità culturale e linguistica è sancito anche all'articolo 3, paragrafo 3 del Trattato sull'Unione europea.

¹⁰ M. CALDERONI, *I postulati della scienza positiva ed il diritto penale*, in G. PAPINI (a cura di), *Scritti di Mario Calderoni*, vol. 1, Società Anon Editrice 'La Voce', Firenze, 1924.

¹¹ U. SCARPELLI, *Filosofia analitica, norme e valori*, cit. p. 16.

¹² *Ibidem*.

¹³ K. SCHWAB, *La quarta rivoluzione industriale*, Milano, Franco Angeli, 2016.

secolo, ma che ha radici in quello precedente. In questo paragrafo si darà conto delle sfumature semantiche di questi termini e delle visioni dell'organizzazione sociale ad esse sottese. Cominciamo con l'individuare i caratteri differenziali del progresso scientifico e tecnologico e il loro ruolo nella qualificazione di questo progresso come rivoluzionario.

Il progresso scientifico e tecnologico ha mostrato nel tempo le due caratteristiche della continuità, intesa come costanza – sebbene con qualche battuta di arresto – e dell'irreversibilità¹⁴, che, a loro volta, ne attestano l'effettività. Possiamo affermare che l'innovazione scientifica e tecnologica si caratterizza per una progressione costante delle conoscenze e che questa progressione è effettiva in quanto se ne possono dimostrare le ricadute sulla qualità della vita di molti individui. Se ciò è vero, è altresì vero che non tutte le tappe di questo progresso possono considerarsi rivoluzionarie. Molte delle straordinarie trasformazioni del passato affermano o estendono tesi, idee e congetture precedenti e non sono rivoluzionarie in senso proprio. Di fatto, si ha una rivoluzione in senso proprio quando cambia il paradigma gnoseologico di riferimento, come accaduto con la c.d. rivoluzione scientifica¹⁵. In altre parole, affinché una scoperta o un avanzamento o un'innovazione possa essere considerata rivoluzionaria, necessita un *quid pluris*, ossia uno stravolgimento sostanziale e profondo dell'usuale modo di pensare, interpretare e rappresentare il reale. Con la rivoluzione scientifica, ad esempio, si è radicalmente trasformata la tradizionale visione finalistica della natura in una visione empirico-descrittiva basata sulla logica della causalità, obbligando così a fissare nuovi criteri per definire la relazione tra l'essere umano e la natura. Diversamente da quanto accade con il progresso dell'IA, nel caso delle rivoluzioni del passato in ambito scientifico, non sono stati gli attori di quella trasformazione radicale a qualificarla come rivoluzionaria. È, infatti, interessante notare che molti protagonisti di eventi che hanno rivoluzionato il sapere, come Copernico e Newton, sebbene fossero rivoluzionari, non si percepivano come tali¹⁶, non solo perché all'epoca il termine non era in uso in campo scientifico, ma soprattutto perché consideravano sé stessi «come epigoni e riscopritori dell'Antichità»¹⁷. Insomma, quello che i posteri definiranno rivoluzionario, veniva percepito da questi pionieri come un ampliamento di idee, di tesi e di teorie formulate da chi li aveva preceduti.

¹⁴N. BOBBIO, *L'età dei diritti*, Torino, Einaudi, 1997, p. 50. Anche Hans Jonas sottolinea che nell'evoluzione del sapere sulla natura «[...] uno sviluppo costante è concepibile non soltanto logicamente, ma storicamente si è verificato fino a oggi anche *de facto*, sia pure con interruzioni, nel modo più manifesto e inconfutabile», cfr. H. JONAS, *Il principio di responsabilità. Un'etica per la civiltà tecnologica*, Torino, Einaudi, 2002, p. 210.

¹⁵N. ABBAGNANO, *Dizionario di Filosofia*, Torino, Utet, 2013, p. 943.

¹⁶*Ibidem*.

¹⁷I.B. COHEN, *La rivoluzione nella scienza*, Milano, Longanesi, 1988, p. 35.

Questo richiamo all'atteggiamento dei grandi scienziati del passato ci consente di effettuare una prima importante distinzione che concerne il significato del termine rivoluzione. Questo termine è di norma utilizzato nel linguaggio ordinario nella sua accezione piena, per indicare un cambio di paradigma concettuale, culturale, sociale ecc. e, in questa accezione, si ricollega a nuovi modelli di indagine del reale, che con quelli antecedenti non hanno più nulla a che vedere. Il termine può essere usato, però, anche in una accezione più ristretta e, in tal senso, fare riferimento al fatto che «nuove scoperte scientifiche vengono qualificate come “rivoluzioni” anche quando non fanno altro che confermare semplicemente e allargare il potere di idee già affermate»¹⁸. Ne consegue che le rivoluzioni scientifiche e tecnologiche in senso proprio sono davvero poche, e che quello che chiamiamo *rivoluzione* molto spesso è solo uno dei passaggi necessari nel processo di miglioramento di aspetti – anche sostanziali – dell'esistenza collettiva e individuale, ovvero non è una vera e propria rivoluzione, bensì un'innovazione, che seppur profonda e importante, non contraddice radicalmente e necessariamente teorie precedenti e non stravolge l'ordine sociale. Porre l'accento su questa sfumatura di significato del termine “rivoluzione” rappresenta il primo passo verso il chiarimento concettuale. Un secondo passo muove nella direzione di evidenziare le criticità legate all'interscambiabilità del termine rivoluzione con quello di innovazione nel linguaggio ordinario.

In tale universo linguistico, questi termini sono spesso utilizzati in maniera interscambiabile¹⁹. In particolare, è l'accezione più ristretta di rivoluzione che viene ad essere oggetto di sinonimia con innovazione. Nel linguaggio ordinario, si usano come sinonimi termini come rivoluzione tecnologica/innovazione tecnologica, rivoluzione digitale/innovazione digitale, rivoluzione algoritmica/innovazione algoritmica e così discorrendo. Questa commistione terminologica può rappresentare un problema in quanto, di norma, nel linguaggio comune il primo significato, che viene in mente, quando si sente il termine rivoluzione, è quello del cambiamento e stravolgimento epocale di determinati modelli e visioni. Il secondo significato di rivoluzione, come interpretazione estensiva di idee e tesi precedentemente elaborate, rimane invece sullo sfondo, ed è però quello che rende questa nozione interscambiabile con la nozione di innovazione. Ecco che allora si insinua il rischio di fraintendimenti circa la portata più o

¹⁸R. LEWONTIN, *Il sogno del genoma umano e altre illusioni della scienza*, Roma-Napoli, Laterza, 2000, p. 35. Sul concetto di rivoluzione cfr. anche H.L. DREYFUS, *What Computers Still Can't Do: A Critique of Artificial Reason*, Cambridge, The MIT Press, (versione originale 1972), 1992, p. 277 ss.

¹⁹In un articolo apparso in *Agenda Digitale* del 2 agosto 2021 a firma di Thomas VACCARI dal titolo *Conosciamo meglio le proteine grazie all'intelligenza artificiale DeepMind: perché è una scoperta enorme*, si parla di come l'IA stia *rivoluzionando il nostro sapere*, riferendosi alla possibilità di incrementare il sapere sul tema oggetto dell'articolo. Ma che questa sia di per sé una rivoluzione nei termini di una rivoluzione scientifica è ancora da dimostrare.

meno radicale delle trasformazioni sociali in atto. Questa può essere definita come una prima trappola linguistica, attraverso la quale si possono influenzare gli umori popolari e non solo.

Se il fenomeno della digitalizzazione della società e del conseguente uso in tutte le attività umane dell'IA viene rappresentato come innovativo nel senso di rivoluzionario, si induce nel pubblico una sensazione di incontrollabilità dei fenomeni che accompagnano i profondi cambiamenti in atto. In pratica, si assume che “accade così e non ci si può fare nulla”, ma non vi è da preoccuparsi perché il cambiamento in atto è positivo. La mescolanza tra il significato di rivoluzione e quello di innovazione consente di celare che i fenomeni di cui parliamo – l'avvento dell'IA, e con essa la trasformazione digitale della società – sono tutt'altro che qualcosa di radicalmente nuovo e inaspettato. Al contrario! Essi fanno parte di un processo in buona parte programmato da diverso tempo e che può essere sia governato e controllato sia regolato nel suo divenire.

La rivoluzione di cui stiamo parlando è allora quella in senso più stretto, che può, a buona ragione, essere considerata come sinonimo di innovazione, le cui caratteristiche sono la continuità e l'irreversibilità. È questo il significato che porta alla luce la possibilità di controllo e di governo del progresso di cui siamo oggi i protagonisti e al contempo i destinatari.

Un terzo passo nell'operazione chiarificatrice è rappresentato dall'analisi dei significati di innovazione. Anche questo termine può essere, infatti, inteso in senso ampio o in senso stretto.

Per innovazione si intende nel linguaggio ordinario «una modificazione che comporta elementi di novità»²⁰ che mirano a migliorare le condizioni di vita delle persone. Tale modificazione può riguardare settori e contesti molto diversi fra loro e chiama in causa forze e strumenti diversificati per le trasformazioni che si vogliono perseguire. Questa definizione generale consente di evidenziare che l'innovazione non è solo quella tecnologica, ma può riguardare altri settori e altri saperi, ad esempio, si parla di innovazione culturale, artistica, sociale ecc.

Nell'attuale momento storico, però, la tendenza prevalente nel discorso istituzionale europeo sull'innovazione è di far coincidere quest'ultima con la sola innovazione tecnologica²¹, lasciando così sullo sfondo la definizione generale di questo termine, che invece include tutte le forme innovative che mirano al miglioramento del benessere dell'umanità.

²⁰ Lo Zingarelli, *Vocabolario della Lingua Italiana*, Bologna, Zanichelli, 2005.

²¹ Su questo tema cfr. A. STIRLING, *Driving Forces in Emerging Technologies: Issues of Social Justice and Democracy in Innovation Governance*, in *Emerging Technologies and Human Rights, Proceedings of the International Conference* organized by the *Committee on Bioethics*, Council of Europe, 2015, pp. 8-33, accessibile al link <https://www.coe.int/en/web/bioethics/emerging-technologies>, ultimo accesso giugno 2022.

In questa operazione di appiattimento semantico del termine innovazione sull'innovazione tecnologica e la qualificazione di quest'ultima come rivoluzionaria, si possono scorgere alcuni messaggi che indirizzano artatamente la costruzione della futura società su modelli prevalentemente, se non esclusivamente, tecnocentrici. Se si assume infatti che l'innovazione sia solo tecnologica ed è questa ad essere rivoluzionaria, si presume che la tecnologia proceda nel suo sviluppo in maniera pressoché autonoma, non sia cioè condizionata da vicende sociali e culturali e sia, quindi, in grado di evolvere attraverso fratture radicali con il passato. A questo modo di intendere l'innovazione tecnologica si accompagna una visione ottimistica, che collega il processo di innovazione tecnologica a cambiamenti sempre e comunque positivi per l'esistenza umana, ossia per tutti gli umani. A sua volta, da questa considerazione discendono decisioni sul piano pratico che toccano gli investimenti nella formazione dei futuri attori della società innovata tecnologicamente. Seguendo questo approccio, infatti, gli investimenti economici andranno prevalentemente, se non esclusivamente, a favore di attività creative e produttive direttamente ed esplicitamente connesse con lo sviluppo tecnologico, vale a dire quello sviluppo tecnologico in grado di rivoluzionare le condizioni della vita umana in meglio.

Sebbene il linguaggio non sia l'unico strumento attraverso il quale si convogliano modelli e visioni socioculturali, è però vero che, nella narrazione sull'IA, la costante commistione tra i termini innovazione e rivoluzione e la riduzione del significato di innovazione a quella tecnologica inducono un atteggiamento di particolare favore per i saperi tecnico-scientifici, attori principali di questi rivolgimenti. Il rischio di un'esasperazione del ruolo di questi saperi è quello di perdere di vista che, per governare gli aspetti innovativi del progresso tecnologico, necessitano abilità e capacità critiche che provengono da altri saperi, segnatamente quelli umanistici.

Il problema non è di poco conto se si pensa che la circoscrizione del significato di innovazione alla sola innovazione tecnologica ha già prodotto scelte politiche, ad esempio, in materia di formazione ed educazione, che puntano a tagliare i costi di tutto ciò che non prepara futuri tecnologi e tecnici²². Ovviamente, nulla in contrario e nessuna valutazione negativa in merito alla preparazione e formazione di queste figure per il contesto lavorativo, il rilievo critico riguarda piuttosto la tendenza, non sempre esplicitata, ma progressivamente più marcata – e già da tempo denunciata²³ – ad accantonare quelle prospettive disciplinari

²²Nel 2021 il Governo inglese ha deciso di incentivare le discipline scientifiche e tecniche a scapito di quelle umanistiche, si veda l'articolo di Adalgisa Marrocco pubblicato sull'*Huffpost* dal titolo *Creano poco lavoro e tanti debiti. Le facoltà umanistiche inglesi chiudono corsi e dipartimenti*, accessibile al link https://www.huffingtonpost.it/entry/creano-poco-lavoro-e-tanti-debiti-le-facolta-umanistiche-inglesi-chiudono-corsi-e-dipartimenti_it_60a7a8dae4b0a24c4f7b33de, ultimo accesso luglio 2021.

²³M. NUSSBAUM, *Non per profitto. Perché le democrazie hanno bisogno della cultura umanistica*, Bologna, il Mulino, 2011.

utili a sviluppare un pensiero critico (*Critical Thinking*). Il pensiero critico consente la giusta distanza dagli accadimenti del presente e di sviluppare una visuale, ad ampio spettro e sul lungo periodo, delle scelte che si operano, invece spesso, sulla scia del contingente interesse economico e di profitto.

Riassumendo le considerazioni sin qui svolte, si può sostenere che l'interscambiabilità tra il concetto di innovazione, intesa solo come progresso tecnologico, e quello di rivoluzione nel suo senso più radicale, in particolare quando questa interscambiabilità è avallata dal linguaggio delle istituzioni politiche, porta i destinatari della comunicazione a prendere – spesso inconsapevolmente – una posizione apprezzativa e selettiva nei confronti di un fenomeno, quello dello sviluppo dell'IA, i cui contorni restano ancora molto sfuocati, in termini di ricadute a medio e lungo termine sulla società e sugli individui. Chiarire, invece, che quello che accade oggi con la trasformazione digitale della società non è una rivoluzione in senso proprio, ma piuttosto una profonda innovazione che parte dal dato tecnologico e che è stata in buona parte programmata nel tempo, consente di far emergere i contrasti e i confronti alla base della dialettica degli opposti, che altro non è se non confronto e contrasto tra atteggiamenti e visioni etiche e valoriali. Se si mette in luce che la questione definitoria e semantica relativa ai termini usati nel dibattito sull'IA non si riduce a vani e sterili scontri terminologici, è possibile evidenziare la varietà dei fini perseguibili attraverso il sapere tecnologico alimentato da altri saperi. Prendendo partito per una qualificazione del progresso tecnologico attuale come innovativo, ma non rivoluzionario in senso proprio, si è portati a evidenziare come esso si basi su un cammino costruito nel tempo su scelte ed esplorazioni di possibilità ed ottimizzazioni delle stesse, che hanno incluso prospettive culturali, sociali, etiche e valoriali e ne hanno escluse altre.

In questo cammino innovativo²⁴, in cui oggi si attribuisce il ruolo esclusivo alla tecnologia, si rischia di lasciare indietro fattori, metodologie, apporti culturali, pratiche, che molti sostenitori dell'eguaglianza e della giustizia sociale²⁵ hanno tentato di includere nelle loro proposte per accompagnare il progresso tecnologico.

Siamo, dunque, giunti a sostenere che l'uso del termine rivoluzione nella narrazione dell'IA è strumentale a un certo modo di impostare il percorso dello sviluppo tecnologico e a portare a compimento specifiche visioni etiche della convivenza sociale. Ci si riferisce a una visione della società in cui prevalgono, ad ogni costo, efficienza e profitto. Ed è al raggiungimento di tali obiettivi che è volta tutta la trasformazione della società industriale in *Società dell'informazione*. In questo modello sociale, «i calcolatori gestiscono la memorizza-

²⁴ A. STIRLING, *Driving Forces in Emerging Technologies: issues of social justice and democracy in innovation governance*, cit.

²⁵ Per tutti, A. SEN, *Lo sviluppo è libertà*, Milano, Mondadori, 2014.

zione, organizzazione e ricerca dell'informazione, governano ogni tipo di macchina, gestiscono i flussi di lavoro, aumentano la realtà con oggetti e contenuti virtuali, animano automi fisici e virtuali»²⁶. In virtù di queste caratteristiche si parla di «un nuovo paradigma socio-tecnologico»²⁷ e di una *quarta rivoluzione*²⁸, che non ha tanto a che vedere con la quarta rivoluzione industriale a opera dell'IA, quanto con una radicale modificazione dell'essere umano che ha al centro l'*infosfera*²⁹. Tuttavia, a parere di chi scrive, il termine rivoluzione riferito al processo di informatizzazione e digitalizzazione della società va ulteriormente specificato per evitare di incorrere in quelle trappole linguistiche correlate all'uso di questo termine nel sentire comune. A ben vedere, la rivoluzione di cui parliamo non ha prodotto un rovesciamento del mondo, ossia non presenta quel *quid pluris* che fa di un'innovazione una rivoluzione in senso proprio. Essa sembra molto più realisticamente il compimento di un processo in atto e pianificato in buona parte da diverso tempo. Sottolineare che questa rivoluzione non è un accadimento inaspettato da cui gli esseri umani sono stati travolti e si trovano oggi a gestire senza precedenti punti di riferimento, ma si tratta piuttosto di un divenire, consente di allontanare la percezione di questo fenomeno come qualcosa che è governato da forze eterodirette che gli individui non possono fare altro che subire. Inoltre, porre l'accento sul fatto che la rivoluzione di cui si parla non è un punto di rottura, ma un processo innovativo *in fieri*, consente di mettere in luce un *bias* cognitivo, che caratterizza spesso il ragionamento umano, limitando la nostra capacità di giudicare e discernere e di operare scelte senza farci eccessivamente sviare nella valutazione delle diverse alternative a disposizione. Il *bias* in oggetto è il c.d. *framing effect*, ossia «il fenomeno per cui se uno stesso problema decisionale viene presentato in maniere differenti, questo condurrà a compiere scelte differenti»,³⁰ perché quello che conta è come le informazioni vengono presentate e non il loro significato informativo. Questo *bias* è strettamente collegato all'uso del linguaggio e alla scelta di presentare con termini che fanno leva sugli aspetti emotivi (appello alle emozioni) il messaggio centrale del proprio argomentare. Se la trasformazione sociale dovuta all'avvento delle tecnologie dell'informazione e dell'IA è presentata come una rivoluzione, intesa come rottura o rovesciamento con il passato e non come un processo innovativo programmato, che si sta portando a compimento, allora sarà più agevole far passare l'idea che quello

²⁶ G. SARTOR, *L'informatica giuridica e le tecnologie dell'informazione. Corso d'informatica giuridica*, Torino, Giappichelli, 2016, p. 1.

²⁷ M. CASTELLS, *The rise of the Network Society*, Oxford, Oxford University Press, 2000, p. 73.

²⁸ L. FLORIDI, *La quarta rivoluzione. Come l'infosfera sta trasformando il mondo*, Milano, Raffaello Cortina, 2017.

²⁹ *Ibidem*.

³⁰ D. CANALE, R. CIUNI, A. FRIGERIO, G. TUZET (a cura di), *Critical Thinking. Un'Introduzione*, Milano, Egea, 2021, p. 9.

che sta accadendo è qualcosa di inevitabile³¹, ovvero corrisponde a un destino a cui non ci si può sottrarre, ma, peggio ancora, a cui non si possono dare delle regole di indirizzo, che eventualmente pongano anche dei limiti. In realtà, il fenomeno a cui stiamo assistendo ha non solo origini storiche e ha seguito un percorso delineato da scelte umane, ma è anche un fenomeno che può essere compreso solo se si abbandona la visione di una tecnologia neutrale, che nasce e si sviluppa con il solo nobile fine di fare del bene e migliorare le vite degli esseri umani. Non è possibile comprendere appieno lo sviluppo tecnologico passato e attuale, unitamente ai corollari filosofico-politici che lo accompagnano, se non viene calato nei meccanismi economici e commerciali che governano le relazioni sociali³². Per dirla con Shoshana Zuboff «[n]on possiamo valutare la strada intrapresa dalla civiltà dell'informazione senza capire bene che la tecnologia non è e non può essere una cosa a sé, isolata da economia e società»³³. Anche se riteniamo che il fenomeno in atto sia un processo innovativo e non una rivoluzione nel senso radicale del termine, ciò non esclude la necessità di adattare tramite interpretazione o elaborare *ex novo* categorie concettuali, soprattutto in ambito giuridico³⁴, per affrontare alcune inedite sfide correlate all'avvento dell'IA. A titolo esemplificativo, Zuboff, nell'analizzare il c.d. capitalismo della sorveglianza, sottolinea che il suo trionfo è dovuto al fatto di *essere senza precedenti* e «[q]uando ci troviamo di fronte a qualcosa senza precedenti, lo interpretiamo automaticamente ricorrendo a categorie familiari, e in tal modo rendiamo del tutto invisibili proprio le sue caratteristiche inedite»³⁵. Per orientare l'avvento dell'IA in direzioni che svelino l'*ideologia inevitabilista*³⁶, dal punto di vista semantico e concettuale risulta indispensabile, per un verso, controllare che il bagaglio di concetti e categorie etiche e giuridiche di cui disponiamo siano ancora utilizzabili nella loro attuale versione e, laddove non lo siano, va operata una trasparente e chiara ricostruzione o una elaborazione *ex novo* al fine di «aggiustare e migliorare il linguaggio, si da

³¹ Cfr. K. KELLY, *L'inevitabile. Le tendenze tecnologiche che rivoluzioneranno il nostro futuro*, Milano, Il Saggiatore, 2017.

³² Cfr. P. SEVERINO (a cura di), *Intelligenza artificiale. Politica, economia, diritto, tecnologia*, Roma, LUISS University Press, 2022.

³³ S. ZUBOFF, *Il capitalismo della sorveglianza. Il futuro dell'umanità nell'era dei nuovi poteri*, Roma, LUISS University Press, 2019, p. 25. Sul tema si veda anche E. SADIN, *Critica della ragione artificiale. Una difesa dell'umanità*, Roma, LUISS University Press, 2019, in particolare il capitolo primo.

³⁴ Cfr. P. BORSELLINO, *Storicità del diritto e filosofia di orientamento analitico-linguistico. Quale rapporto?*, in A. BALLARINI (a cura di), *La storicità del diritto. Esistenza materiale, filosofia, ermeneutica*, Torino, Giappichelli, 2018, pp. 107-128.

³⁵ S. ZUBOFF, *Il capitalismo della sorveglianza*, cit. p. 22.

³⁶ *Ibidem*, p. 236.

rifarne un buon mezzo per l'uomo con i suoi attuali atteggiamenti, bisogni e progetti, nelle circostanze in cui si trova»³⁷.

Un'ulteriore ragione che si può addurre a favore dell'interpretazione del processo di trasformazione digitale e informatizzata della società come compimento di un percorso in buona parte programmato, piuttosto che come rottura con il passato, è il suo inquadramento entro una tendenza filosofica della tradizione culturale occidentale, volta spasmodicamente a eliminare incertezze e garantire controllo³⁸: una tendenza che alimenta e si autoalimenta di una retorica oggettivistica che definisce il reale, inclusi i comportamenti umani, solo in termini di strutture di dati intellegibili, autoevidenti e ridicibili a una spiegazione chiara e semplice³⁹. Se si tiene conto di questa tendenza, che Dreyfus inquadra utilizzando l'etichetta di *ontological assumption*⁴⁰, si può avanzare l'ipotesi che un'eccessiva enfasi posta sulla portata rivoluzionaria, intesa nel suo significato più drastico e drammatico, può apparire ideologica⁴¹, perché nasconde ai destinatari della trasformazione l'origine valoriale e pianificata della stessa⁴².

Una delle principali difficoltà che incontra il grande pubblico, a cui si rivolge la narrazione dell'IA improntata alla dialettica degli opposti, è quella di non riuscire a cogliere i riflessi di visioni riduzionistiche e oggettivizzanti dell'umano, sottesi a tale prospettazione dialettica. Questo può avere ricadute gravi sull'atteggiamento individuale e collettivo. Di fatto, sono queste visioni diffuse e penetrate nella cultura che spesso contribuiscono all'indifferentismo e alla rinuncia di chi si sente in balia del fato, perché inibiscono gli stimoli e le ragioni per mantenere atteggiamenti proattivi verso gli accadimenti della vita. La dialettica

³⁷ U. SCARPELLI, *Filosofia analitica, norme e valori*, cit. p. 30.

³⁸ Una tendenza che, attraverso le tecnologie intelligenti, si sta portando a compimento, ma che viene criticata, ad esempio, da Laurent Alexandre, il quale sostiene che va salvaguardata una componente di casualità nell'agire umano a tutela della nostra stessa salute mentale. «Nel gioco della vita, dobbiamo assolutamente preservare una parte di casualità, altrimenti, molto semplicemente, non c'è più partita», L. ALEXANDRE, *La guerra delle intelligenze. Intelligenza artificiale contro intelligenza umana*, Torino, EDT, 2018, p. 257.

³⁹ H.L. DREYFUS, *What Computers Still Can't Do: A Critique of Artificial Reason*, cit. p. 211 ss. L'autore sottolinea che «[...] the goal of the philosophical tradition embedded in our culture is to eliminate uncertainty: moral, intellectual, and practical». Cfr. anche A. ROUVROY, B. STIEGLER, *Il regime di verità digitale. Dalla governamentalità algoritmica a un nuovo Stato di diritto*, in *La Deleuziana-Rivista Online di Filosofia*, 2016, 3, pp. 6-28.

⁴⁰ H.L. DREYFUS, *What Computers Still Can't Do: A Critique of Artificial Reason*, cit. p. 206 ss.

⁴¹ Nei sensi in cui, ad esempio, intende Kelsen questo termine. H. KELSEN, *Lineamenti di dottrina pura del diritto*, Torino, Einaudi, 1952, p. 60. Secondo l'autore «[l'] ideologia [...] nasconde la realtà perché, o la esalta con l'intento di conservarla o di difenderla, oppure la deforma con l'intenzione di distaccarla, di distruggerla e di sostituirla con un'altra».

⁴² E. SADIN, *Critica della ragione artificiale*, cit. p. 38. L'autore ritiene che stiamo entrando nell'era antropomorfa della tecnica a cui sarebbe sotteso il desiderio di «instaurazione di un'organizzazione più affidabile e perfetta delle cose».

degli opposti, alimentata dall'antropomorfismo preconconcetto, gioca un ruolo determinante nella tendenza alla «schematizzazione univoca e definitiva»⁴³ dell'esperienza umana in quanto non lascia spazio né al linguaggio della libertà né a un'etica della responsabilità.

Parlare di innovazione come rivoluzione, lasciando intendere il senso drastico di questo termine, oppure chiarire che la rivoluzione di cui si fa pubblicità è solo il punto di arrivo di un percorso, cambia molto la prospettiva di analisi e di percezione dei fenomeni attuali, perché, nel primo caso, la strada sembra in qualche modo già tracciata, mentre nel secondo caso, sembra tracciabile attraverso l'apporto di regole volte a bilanciare i diversi interessi in gioco. Detto diversamente, se il progresso tecnologico viene equiparato a un accadimento improvviso e travolgente come una rivoluzione, il messaggio che passa è che non esiste più nulla che si possa fare *a priori* per individuare i fini che con il progresso tecnologico si intendono perseguire: essi sono già predefiniti. In tale scenario, vengono meno risposte, perché perdono senso alcune fondamentali domande. Non ci chiediamo tanto perché questo sviluppo accada e sia così fortemente accelerato, perché non ci interessa e non possiamo sapere perché ciò avviene. Ci accontentiamo della presa d'atto che avviene e basta. Questa è la filosofia sottesa agli atteggiamenti dei grandi attori che guidano il progresso tecnologico perché detengono le tecnologie per farlo. Per dare conto di ciò, basta ricordare che nel 2008 Anderson scrive «Google's founding philosophy is that we don't know why this page is better than that one: If the statistics of incoming links say it is, that's good enough. No semantic or causal analysis is required»⁴⁴. Non interessano interrogativi sulle cause e sulle relazioni semantiche di questo progresso, ma solo che sta avvenendo. Il punto è che siamo spinti a non indagare le ragioni giustificanti di questo progresso e, quindi, anche a non cercare argomenti con cui affrontarlo criticamente. Come ha osservato Shoshana Zuboff abbiamo perso il controllo di domande cruciali di cui, invece, altri hanno preso il controllo. Quali domande? Quelle che «definiscono la conoscenza, l'autorità e il potere nella nostra epoca: *Chi sa? Chi decide? Chi decide chi decide?*»⁴⁵.

Questa c.d. rivoluzione digitale e/o algoritmica lungi dall'essere una vera rivoluzione, non è in realtà null'altro che il compimento di un lungo percorso che ha avuto nel secolo XX un'accelerazione decisiva. Un percorso tracciato da chi deteneva e detiene la tecnologia e il potere economico⁴⁶ e che ha potuto impor-

⁴³ *Ibidem*, p. 29.

⁴⁴ C. ANDERSON, *The End of Theory: The Data Deluge Makes the Scientific Method Obsolete*, in *Wired*, 23 giugno 2008, <https://www.wired.com/2008/06/pb-theory/>, ultimo accesso luglio 2021.

⁴⁵ S. ZUBOFF, *Il capitalismo della sorveglianza*, cit. p. 187.

⁴⁶ Per un riassunto della storia dell'IA cfr. N. BOSTROM, *Superintelligenza. Tendenze, pericoli, strategie*, Torino, Bollati-Boringhieri, 2018; M. IENCA, *Intelligenza2. Per un'unione di intelli-*

re, in maniera non democratica⁴⁷, un certo indirizzo. Il processo ha avuto momenti di stasi⁴⁸ in passato, anche perché non poche sono state le difficoltà incontrate nell'implementazione della visione digitalizzata della società. Si pensi a titolo esemplificativo alla digitalizzazione della Pubblica Amministrazione, che ha incontrato e ancora incontra sia problemi pratici di trasferimento del dato analogico su base digitale sia problemi etico-giuridici relativi alla realizzazione della c.d. trasparenza⁴⁹.

Nonostante i momenti di apparente arresto, il processo è proseguito e da un certo momento in poi ha seguito un passo celere. Con la pandemia dovuta al COVID-19, l'accelerazione è stata fortissima. L'umanità si è ritrovata catapultata dentro una dimensione che ancora non le apparteneva globalmente e questa dimensione digitale, sebbene ci abbia salvato sotto tanti profili, preservando, ad esempio, l'attività lavorativa per molti, ha mostrato però anche tutti i suoi limiti, ossia i limiti di una forzatura delle relazioni umane spinte alla totale virtualizzazione sotto il controllo dell'occhio digitale⁵⁰. In altre parole, è stata garantita certamente una *participatio* alle attività che, grazie all'ausilio tecnologico, è possibile svolgere. Tuttavia, questa partecipazione è avvenuta per molti senza la necessaria consapevolezza delle caratteristiche delle tecnologie usate e della loro pervasività.

In pratica, si è reso palese lo scarto enorme tra quello che la tecnologia ha reso possibile fare in poco tempo, a fronte dell'emergenza, e la preparazione critica delle persone a gestire questa tecnologia⁵¹ per periodi anche molto lunghi. Ma questo scarto non dipende dal fatto che il processo a cui assistiamo sia di

genza naturale e artificiale, Torino, Rosenberg & Sellier, 2019; L. ALEXANDRE, *La guerra delle intelligenze*, cit.

⁴⁷ G. FIORIGLIO, *Democrazia elettronica. Presupposti e strumenti*, Milano, Wolters Kluwer-Cedam, 2017. Come ricorda l'autore «[i]l fatto che l'utilizzo delle nuove tecnologie sia in crescita presso tutte le fasce delle varie popolazioni non deve tuttavia portare alla frettolosa considerazione per cui le tecnologie dell'informazione e della comunicazione sarebbero di per sé neutrali e democratiche: anzi esse possono fungere da Cavallo di Troia per i poteri pubblici e privati, per cui le tecnologie risultano neutrali e sicure solo in apparenza così da celare finalità ulteriori come la sorveglianza dei loro possessori, la loro profilazione e la personalizzazione dei flussi informativi», p. 27.

⁴⁸ Per quello che riguarda il progresso nel campo dell'IA, si usa parlare di *primavera* e *inverni*, cfr. N. BOSTROM, *Superintelligenza. Tendenze, pericoli, strategie*, cit.

⁴⁹ Cfr. S. ROSSA, *Trasparenza e accesso all'epoca dell'amministrazione digitale*, in R. CAVALLO PERIN, D.U. GALETTA (a cura di), *Il diritto dell'amministrazione pubblica digitale*, Torino, Giappichelli, 2020, pp. 247-275. Torneremo sul tema della trasparenza nei capitoli successivi.

⁵⁰ In un articolo comparso sul *Corriere della Sera*, Paolo Benanti e Sebastiano Maffettone sottolineano che «per quel che ne sappiamo la formazione del soggetto è sociale e dipende evolutivamente anche dall'interazione effettiva tra persone», *Algoritmi e Nuovi Strumenti di Partecipazione Democratica*, 27 luglio 2021, da p. 24.

⁵¹ Sul tema dell'informazione ed educazione all'uso responsabile del web cfr. L. PALAZZANI, *Tecnologie dell'Informazione e Intelligenza Artificiale. Sfide Etiche al Diritto*, Roma, Studium edizioni, 2020.

per sé rivoluzionario, quanto dal fatto che nella programmazione del progresso tecnologico, partito nel secolo scorso, non si è incluso *ab initio*, e in maniera altrettanto pianificata, l'innalzamento degli standard di conoscenza, di sensibilizzazione e di intervento propriamente educativo-pedagogico nei confronti di tutti i destinatari⁵². Questo non deve sorprendere.

Da quando il mondo ha cominciato a essere interconnesso, prima con Internet⁵³ e successivamente con il cellulare e infine con i social media, in generale l'atteggiamento di chi possiede la tecnologia e la forza economica per promuoverla è stato quello di mettere a disposizione un crescente numero di strumenti per un pubblico sempre più ampio, ma non quello di costruire strutturalmente e sistematicamente, attorno a tali strumenti, anche interventi volti a sviluppare consapevolezza delle reali caratteristiche di ciò che si utilizza⁵⁴. Si potrebbe sostenere che questo non è compito dei privati, ma della sfera pubblica, che dovrebbe indicare principi e valori per governare il progresso tecnologico. Se ciò è vero, è altresì vero che le grandi multinazionali non possono considerarsi privati pari agli altri. Esse hanno potuto esercitare il potere di ritardare l'intervento pubblico, in particolare attraverso la regolazione giuridica, di molte delle attività da loro svolte. Questo è, ad esempio, ben documentato da Shoshana Zuboff quando richiama il motto della Silicon Valley: *innovazione senza permesso*⁵⁵. L'intervento normativo istituzionale è stato fortemente contrastato per poter affermare con forza che l'autoregolamentazione tramite codici di condotta fosse la via da perseguire. Questa scelta cela ovviamente il più grande *bias* dell'autoregolamentazione che avviene senza un quadro giuridico definito, ossia che fa solo

⁵² Il problema, ad esempio, delle fasce di popolazione anziana e del loro rapporto con le tecnologie è ormai al centro del dibattito europeo. Ma poteva essere affrontato in modo sistematico in tempi più risalenti, quando già si poteva prevedere il rischio dell'esclusione di intere fasce di età. Oggi se ne dà conto ne *Il Libro verde sull'invecchiamento demografico* della Commissione europea del 2021 che prende in considerazione anche il problema del rapporto tra la popolazione anziana e la digitalizzazione delle attività, accessibile al link https://ec.europa.eu/info/sites/default/files/com_2021_50_f1_green_paper_it.pdf, ultimo accesso luglio 2021.

⁵³ La pervasività e globalità di Internet ha avuto come conseguenza che «[n]on solo duplichiamo nel ciberspazio attività che siamo soliti compiere nello spazio fisico, ma Internet sta modificando le modalità nelle quali si svolge ogni tipo di attività, dalla ricerca scientifica, alla produzione, alla socializzazione» cfr. G. SARTOR, *L'informatica giuridica e le tecnologie dell'informazione*, cit. p. 207.

⁵⁴ M.G. LOSANO, *La lunga marcia dell'informatica nelle istituzioni italiane*, in R. CAVALLO PERIN, D.U. GALETTA (a cura di), *Il diritto dell'amministrazione pubblica digitale*, Torino, Giappichelli, 2020, pp. XXIII-LVIII. L'autore nel tracciare brevemente il dibattito sull'informatizzazione della Pubblica Amministrazione italiana chiarisce che «[...] per raggiungere questa digitalizzazione non basta comperare tante attrezzature informatiche: bisogna anche saperle usare, e proprio qui si manifesta il vero problema dell'Italia. Da un lato, bisogna ristrutturare la gestione sia pubblica sia privata adattandola all'informatica; dall'altro, bisogna preparare le persone all'uso intelligente degli strumenti informatici e, a questo fine, bisogna ripensare e inventare alcuni percorsi didattici negli istituti tecnici e nelle università», p. XXIV.

⁵⁵ S. ZUBOFF, *Il capitalismo della sorveglianza*, cit. p. 60.

gli interessi di chi si autoregola. Questa scelta è profondamente antidemocratica e anti egualitaria. Non si nega che in un mondo molto complesso come quello dello sviluppo tecnologico siano di aiuto diverse fonti di regolazione, ma esse devono collocarsi entro una cornice definita da chi deve garantire il processo democratico ed egualitario di queste trasformazioni.

Ecco perché sarebbe fruttuosa una cooperazione seria tra gli attori privati e quelli pubblici per elevare le capacità di interazione e utilizzo degli utenti, superando la visione dualistica tra privato e pubblico. Con ciò non si vuole proporre una privatizzazione del governo delle nuove tecnologie. Tutt'altro! Quello che si intende dire è che la ricezione dei principi e delle regole elaborate a livello istituzionale deve trovare un alleato nell'implementazione sul campo da parte dei privati a cui queste regole si rivolgono. In questo modo sarà possibile incentivare una visione comune della società del futuro in cui siano chiare le soglie incompatibili con il rispetto della persona. Già Stefano Rodotà si esprimeva in questa direzione, ma nonostante siano passati ormai diversi anni dalle sue riflessioni⁵⁶, siamo ancora in una situazione in cui la cooperazione tra pubblico e privato su questi temi è insoddisfacente.

Nelle pieghe dei complessi rapporti tra pubblico e privato si è insinuata la dialettica degli opposti, che confonde i piani⁵⁷ ed estremizza le tensioni. La narrazione antropomorfizzata, che accompagna la creazione di artefatti da parte dell'uomo, rende difficile stabilire la giusta distanza tra noi e le nostre invenzioni⁵⁸.

⁵⁶ S. RODOTÀ, *Il diritto di avere diritti*, Roma-Napoli, Laterza, 2012.

⁵⁷ Come osserva J. TESTART, A. ROUSSEAU, *Au péril de l'humain. Les promesses suicidaires des transhumanistes*, Paris, Éditions du Seuil, 2018, p. 149, «[m]ais il est absurde de dire que l'homme a été battu par l'intelligence artificielle: ce son bien des humaines qui ont conçu cette machine qui gagne, fait remarquer le philosophe Jean-Jacques Delfour. Il n'y a pas "d'intelligence machine", puisque celle-ci est une invention humaine».

⁵⁸ Molto eloquente è la vicenda salita agli onori della cronaca a giugno 2022. Blake Lemoine, ex ingegnere della Google, licenziato dall'azienda, ha pubblicato il presunto testo dell'intervista che ha fatto a uno dei chatbot di IA più sofisticati dell'azienda, denominato *Language Model for Dialogue Applications* (LaMda). In tale intervista LaMda sostiene di essere una *persona* e quindi di essere senziente, provare emozioni ed avere esperienze soggettive esattamente come gli esseri umani. La notizia ha suscitato scalpore e preoccupazione. Ma ancora una volta il problema è in buona parte legato al linguaggio e alla selezione delle parole che l'IA può effettuare attingendo alle fonti internet per dare l'impressione di essere ciò che non è. L'articolo pubblicato su *WIRED.it* del 15 giugno 2022, a firma di Katherine CROSS, dal titolo *Sulla coscienza delle intelligenze artificiali ci stiamo facendo le domande sbagliate*, centra il cuore del problema quando afferma che «invece di chiederci se possono diventare senzienti, dovremmo concentrarci sui possibili rischi che comporta il loro eccessivo realismo», aggiungendo che «il rischio è rappresentato dal fatto che siamo disposti a creare macchine sofisticate in grado di imitare gli esseri umani al punto tale che non possiamo fare a meno di antropomorfizzarle, oltre che da come i colossi tecnologici sono in grado di sfruttare questo aspetto in modi profondamente immorali». Il testo dell'articolo è accessibile a questo link <https://www.wired.it/article/intelligenza-artificiale-senziente-lamda-rischi/>, ultimo accesso giugno 2022. Cfr. anche l'articolo di Khari JOHNSON del 18 giugno 2022 da titolo *LaMda e la*

Se, sul piano istituzionale e politico, vi sono state drammatiche mancanze in merito a tempestive iniziative⁵⁹ volte a «fornire servizi educativi che non siano unicamente finalizzati all'apprendimento delle modalità di utilizzo di determinati strumenti informatici, ma altresì alle metodologie che consentono un approccio critico alle medesime»⁶⁰, sul piano della comunicazione mediatica, invece, la narrazione dell'avvento dell'IA ha proseguito lungo la strada tracciata dalla cinematografia e dalla letteratura in materia⁶¹, attraverso una tassonomia concettuale antropomorfizzante che consente di consolidare quella prospettiva dialettica. Da questa narrazione e dall'ubiquità dell'uso di molte nozioni, in quanto presenti sia nel linguaggio ordinario sia in quello tecnico e giuridico, derivano molti dei problemi di comprensione della reale portata di quello che sta accadendo. Se, come abbiamo cercato di dimostrare in questo paragrafo, il fenomeno tecnologico attuale è una profonda innovazione per buona parte programmata e non una rivoluzione in senso proprio, anche i termini centrali della dialettica degli opposti, ossia responsabilità, persona, autonomia e quelli a essi connessi, debbono essere ridefiniti per far emergere il preponderante ruolo e la centralità dell'uomo nel delineare i contorni di questa innovazione. Tale operazione di controllo sui significati sarà, invero, utile per abbandonare sostanzialismi inficiati da ipostatizzazioni sottesi alla dialettica degli opposti.

Dovendo, quindi, soffermarci sulla questione definitoria, dedicheremo un po' di spazio a questo tema nel prossimo paragrafo, richiamando alcuni elementi centrali della teoria della definizione e delle funzioni del linguaggio.

3. Diffinitio quid nominis *v.* diffinitio quid rei

Per proporre l'analisi di un fenomeno attraverso l'indagine dei concetti che ne costituiscono i principali elementi di narrazione, occorre contestualizzarla entro i confini delle funzioni del linguaggio che fungono da presupposto per la questione definitoria.

trappola dell'intelligenza artificiale senziente, in *Wired.it*, <https://www.wired.it/article/intelligenza-artificiale-lambda-dibattito-at-senziente/>.

⁵⁹ Consideriamo a titolo esemplificativo la Raccomandazione del Consiglio d'Europa (2019) 10 relativa allo sviluppo e promozione della cittadinanza digitale nella quale si torna sia sulla questione del diritto di accesso sia sul problema informativo e della conoscenza cfr. https://search.coe.int/cm/Pages/result_details.aspx?OBJECTID=090000168098de08, ultimo accesso agosto 2021.

⁶⁰ G. FIORIGLIO, *Democrazia elettronica. Presupposti e strumenti*, cit. p. 215.

⁶¹ Come testimoniato dai titoli degli articoli di giornale in materia. Alcuni esempi: J. D'ALESSANDRO, *L'educazione delle macchine*, in *La Repubblica*, 18 ottobre 2017; R. MANTOVANI, *Eva il primo robot che ricambia il sorriso degli umani*, in *Gazzetta del Sud*, 4 giugno 2021; E. ZACHEO, *La realtà degli umani negli occhi di un robot*, in *Nuovo Quotidiano di Puglia*, 28 giugno 2021.

Questa contestualizzazione permette di evidenziare le trappole linguistiche che alimentano l'ideologia antropomorfizzante alla base della dialettica degli opposti, in quanto tali trappole sono costruite sia attorno al significato ambiguo e vago dei termini impiegati sia alla mescolanza delle funzioni del linguaggio.

Attraverso questa commistione funzionale si possono confondere i piani del controllo sul comportamento umano che il linguaggio aiuta a svolgere.

Come rilevato, invero a suo tempo, dall'orientamento filosofico pragmatista «in ogni suo uso il linguaggio ha una funzione di guida del comportamento umano»⁶². Vanno, però, distinte la funzione di guida diretta e quella di guida indiretta. La differenza tra le due dipende dal diverso uso che si fa del linguaggio in funzione prescrittiva o descrittiva. Se la prima funzione è quella che porta un soggetto a compiere direttamente l'azione prescritta, la funzione assertivo-descrittiva o conoscitiva funge, invece, da stimolo indiretto al compimento di azioni. La funzione prescrittiva è svolta da proposizioni che prescrivono comportamenti avvalendosi di norme, principi, valori – morali, giuridici, sociali – e chiama in causa la sfera del normativo. Le prescrizioni promuovono direttamente, attraverso ordini, obblighi, doveri, sanzioni e/o incentivi, i comportamenti desiderati. Questa funzione si rivela attraverso giudizi di opportunità o inopportunità, appropriatezza o inappropriata, giustizia o ingiustizia circa le ricadute dei comportamenti prescritti.

La funzione descrittiva o conoscitiva, invece, procede mediante asserzioni che vertono su fatti. Essa è il dominio del vero e del falso, ossia di una verità determinabile mediante analisi logica degli enunciati e/o controllo empirico. In questo ambito si inseriscono le asserzioni della scienza.

Attraverso le asserzioni di cui si avvale il linguaggio in funzione conoscitiva si possono stimolare indirettamente comportamenti. Invero, aprendo, mediante le scoperte scientifiche e le loro conversioni tecniche, nuovi scenari di azione si allargano gli orizzonti di scelta degli individui. In altre parole, gli individui si trovano di fronte a inedite possibilità di intervento, in virtù della disponibilità di strumenti sino a quel momento sconosciuti⁶³. Si pensi, a titolo esemplificativo, alle tecniche di riproduzione assistita. Questo importante traguardo ha trasformato profondamente l'ambito riproduttivo, consentendo la riproduzione senza rapporto sessuale. Un altro esempio, più vicino alla tematica di questo volume, è l'avvento di Internet, che ha consentito la comunicazione a livello globale per tutti, allontanando il problema della distanza fisica.

Le possibilità di azione messe a disposizione dalla scienza e dalla tecnologia non si limitano ad avere ricadute sul piano delle opportunità in chiave utilitaristica. Sottesi ai nuovi scenari vi sono scelte di valore che obbligano a una rifles-

⁶² U. SCARPELLI, *Definizione nel diritto*, in ID. (a cura di), *L'etica senza verità*, Bologna, Il Mulino, 1982, p. 207.

⁶³ U. SCARPELLI, *Scienza, Sapere, Sapienza*, in *Rivista Internazionale di Filosofia del Diritto*, 1986, IV Serie, LXIII, pp. 245-258.

sione sulla sostenibilità etica delle nuove opportunità e sulla necessità o meno di prevedere delle specifiche condizioni di realizzazione.

Quando si parla di scelte di valore bisogna intendersi. Non si deve, infatti, incorrere nell'errore di ritenere che l'ambito descrittivo e quello normativo rappresentino una sorta di linea continua, dove le prescrizioni sono derivabili dalle descrizioni. Questo tipo di conclusione è stato oggetto di profonde critiche nell'ambito del dibattito attorno alla c.d. *is-ought question*, ossia alla tesi della *Grande Divisione*. Il dibattito attorno allo sfondo ontologico della distinzione tra funzioni del linguaggio è stato al centro di vivaci scontri, soprattutto nel secolo XX. Come ci ricorda Uberto Scarpelli, per alcuni la distinzione corrisponde «a una scissione della realtà, e dell'esperienza che ne hanno gli uomini, in due dimensioni, la dimensione dell'essere e la dimensione del dover essere», mentre per altri siamo di fronte a due diversi usi del linguaggio «nel contesto unico e continuo dell'esperienza umana»⁶⁴. Trattasi di un dibattito non fine a sé stesso o di interesse solo accademico. Anzi! Vi sono importanti ricadute pratiche sulla trasparenza delle posizioni etiche sottese.

Di fatto, aderire alla tesi della Grande Divisione tra descrittivo e normativo, vuol dire prendere partito per una trasparenza discorsiva, che prende le mosse dal riconoscimento che non è possibile transitare dalla funzione descrittiva del linguaggio a quella prescrittiva con mezzi puramente logico-linguistici, a meno di non compiere un salto logico⁶⁵. Chi sostiene questa posizione ritiene possibile costruire universi di discorso trasparenti, in quanto suscettibili di rigore argomentativo sottoponibile a controllo intersoggettivo, senza confondere questo metodo con il metodo sperimentale proprio del contesto strettamente scientifico.

Il confronto sulla *is-ought question*, nelle sue svariate sfumature, non è del tutto risolto, perché non possiamo sostenere che si sia definitivamente raggiunto un accordo condiviso circa la portata della Grande Divisione tra le due funzioni del linguaggio. Quello che però pare certamente acquisito è il riconoscimento che mantenere la distinzione tra le funzioni del linguaggio corrisponde meglio all'esigenza di discorsi trasparenti rispetto alle premesse e alle conclusioni da queste derivate tramite il controllo dei passaggi intermedi, ossia un controllo logico e/o empirico.

Va da sé allora che, anche se lo scontro tra divisionisti e antidivisionisti è ancora aperto, sono chiare le conseguenze dell'appartenenza all'uno o all'altro schieramento. Mantenere saldo questo punto consente di affrontare alcune novità critiche dell'attuale sviluppo tecnologico che possono dare, apparentemente, nuova linfa alla risalente disputa.

⁶⁴ U. SCARPELLI, *La «grande divisione» e la filosofia della politica*, in ID., *L'etica senza verità*, Bologna, Il Mulino, 1982, pp. 115-139.

⁶⁵ *Ibidem*, p. 115.

Come avremo modo di analizzare meglio nel prossimo capitolo, vi sono alcuni degli attuali sviluppi tecnologici che mettono in crisi la richiamata distinzione tra influenza diretta o indiretta sui comportamenti umani del progresso scientifico-tecnologico e che, di conseguenza, rendono arduo difendere la separazione tra le due funzioni del linguaggio. Alcune tecnologie riescono, infatti, a incidere direttamente sulle condotte, inibendo o favorendo direttamente l'atteggiamento che risponde a un determinato obiettivo. È questo il caso di quelle tecnologie che consentono un controllo eterodiretto di macchine e strumenti che appartengono a un dato individuo, ad esempio, quando a chi non paga l'assicurazione dell'automobile si può impedire che possa accenderla per usarla⁶⁶. A fronte di questa novità, ci si può chiedere se si possa ritenere definitivamente risolta la *is-ought question* in favore di chi «nega la grande divisione perché con essa si divide troppo» o se vi siano ragioni rilevanti per continuare a difendere questa distinzione. Chi parteggia per questa seconda ipotesi pone l'accento sulla necessità di trasparenza nella prospettazione di tesi e di prese di posizioni rispetto a questioni dalla portata trasformativa molto ampia, come è il caso dell'attuale innovazione tecnologica. Distinguere tra i fatti, le operazioni e i discorsi possibili intorno ai fatti rappresenta una forma di controllo sui fini e i valori sottesi a scelte di indirizzo della trasformazione digitale della società. Questa trasparenza è tanto più richiesta quanto più si presentano situazioni nel linguaggio ordinario e in quello istituzionale attorno alle vicende dell'IA, in cui diventa difficile distinguere tra la funzione conoscitiva e quella prescrittiva. Ciò perché queste due funzioni nella comunicazione non si svolgono sempre in autonomia e non sono sempre facilmente distinguibili. Spesso i termini usati e/o la modalità di combinazione delle parole negli enunciati nasconde una valenza persuasiva che fa leva sugli aspetti emotivi, rendendo difficile per il destinatario distinguere tra le indicazioni informative e quelle prescrittive del messaggio che riceve.

Nella narrazione dell'IA, la commistione tra le diverse funzioni e usi del linguaggio, spesso indotta dalla valenza persuasiva, rappresenta una delle maggiori fonti di problematicità, soprattutto per i destinatari finali della trasformazione digitale, ossia gli utenti-cittadini. Quando si narra dell'IA, non solo mediaticamente, ma anche nel dibattito istituzionale, non è sempre facile distinguere tra la funzione conoscitiva e quella prescrittiva. Diverse sono, invero, le situazioni discorsive, come di seguito esemplificate.

Ci sono situazioni in cui è abbastanza semplice ricondurre il racconto dentro i confini di una delle due funzioni del linguaggio. È questo il caso quando lo scopo è puramente informativo e si vuole raccontare, ad esempio, quali sono i

⁶⁶ Commenteremo questo uso della tecnologia nei prossimi capitoli. Qui anticipiamo una domanda a cui risponderemo, ossia se siamo sicuri che il fatto che l'IA possa impedire o imporre un comportamento bloccando direttamente azioni da parte di un individuo metta davvero in crisi la Grande Divisione.

potenziali impieghi delle tecnologie guidate dall'Intelligenza Artificiale o ricordare come si sono ottenuti i risultati raggiunti⁶⁷. Altrettanto chiaro è il caso in cui se ne scrive utilizzando esplicite modalità prescrittive, per indicare la necessità di percorrere una data via per il suo sviluppo⁶⁸. Ad esempio, quando Ursula von der Leyen, Presidente della Commissione europea, sostiene che «[...] Dobbiamo adesso fare di questo decennio il decennio digitale europeo affinché tutti i cittadini e le imprese possano avere accesso a quanto di meglio può offrire il mondo digitale. La bussola per il digitale presentata oggi ci indica chiaramente la rotta da seguire per realizzare questo obiettivo»⁶⁹.

Ma molto più spesso, la narrazione si carica di toni trionfalistici ed evocativi, di valori “euforici”. In questo modo la valenza persuasiva si mimetizza con la descrizione e/o con la prescrizione. Ne è un esempio la dichiarazione di Margrethe Vestager, Vicepresidente esecutiva per un'Europa pronta per l'era digitale, la quale così si esprime «[a]ssieme al Parlamento europeo, agli Stati membri e agli altri portatori di interessi, lavoreremo affinché l'Europa diventi il partner prospero, determinato e aperto che vogliamo che sia a livello mondiale e per far sì che ciascuno di noi possa beneficiare appieno del benessere generato da una società digitale inclusiva»⁷⁰.

Una volta posta la distinzione tra le funzioni del linguaggio, occorre volgere l'attenzione alla questione definitoria. Tale questione chiama in causa non solo la separazione tra funzione prescrittiva e conoscitiva del linguaggio, ma anche tra diverse concezioni del linguaggio, in particolare tra la concezione convenzionalistica e quella essenzialistica⁷¹. Secondo la prima concezione, il linguaggio è una istituzione culturale, ossia un mezzo per la comunicazione tra i consociati, e si esprime attraverso parole, ossia simboli artificiali, combinati in enunciati

⁶⁷ Ad esempio, la storia dell'IA narrata sul sito del Consiglio d'Europa, <https://www.coe.int/en/web/artificial-intelligence/history-of-ai>: oppure il titolo neutro di *Wired* del 23 giugno 2021 *L'evoluzione della pratica clinica, tra intelligenza artificiale e Real World Data*, di Luca DOTTI accessibile al seguente link <https://www.wired.it/scienza/medicina/2021/06/23/evoluzione-pratica-clinica-intelligenza-artificiale-real-world-data/>, ultimo accesso agosto 2022; sul *Corriere Salute* del 13 maggio 2021 il titolo recita *UE, Intelligenza artificiale e nuove regole*, di Eugenio SANTORO, accessibile al link https://www.corriere.it/salute/ehealth/22_giugno_09/nuove-regole-l-intelligenza-artificiale-campo-clinico-c7f199b2-e288-11ec-9f19-f9603cda965c.shtml, ultimo accesso agosto 2022.

⁶⁸ A titolo esemplificativo, il *Corriere della Sera* del 31 maggio 2021 riportava il titolo *Robot? No, Intelligenza artificiale. Le leggi di Asimov non bastano più* di Giovanni LO STORTO e Daniele MANCA, accessibile al link https://www.corriere.it/cultura/21_maggio_29/intelligenza-artificiale-non-robot-leggi-asimov-non-bastano-piu-266939b6-c08c-11eb-b483-a79329df2c54.shtml, ultimo accesso agosto 2022.

⁶⁹ Testo del discorso pubblicato sul sito della Commissione europea accessibile al seguente link https://ec.europa.eu/commission/presscorner/detail/it/IP_21_983.

⁷⁰ Testo del discorso pubblicato sul sito della Commissione europea accessibile al seguente link https://ec.europa.eu/commission/presscorner/detail/it/IP_21_983.

⁷¹ Cfr. sul tema M. BARBERIS, *Filosofia del diritto*, Torino, Giappichelli, 2008.

aventi un significato. Sul linguaggio, così inteso, è possibile un controllo linguistico attraverso aggiustamenti di significati, che avvengono per il trascorrere del tempo con l'introduzione di nuove parole o con la rivisitazione dei significati dei termini già esistenti alla luce dei cambiamenti sociali e culturali. In questa prospettiva diventa rilevante la determinazione delle condizioni dell'uso di un termine in relazione ai fini che si vogliono perseguire. Nel linguaggio ordinario ciò è dato dalle condizioni di uso consolidatesi tra i parlanti di una determinata comunità linguistica.

Diversamente dall'approccio appena descritto, la concezione del linguaggio in chiave essenzialistica parte da un rapporto speculare tra realtà e parole. In altri termini, tra le parole e ciò che significano intercorrerebbe una relazione *naturale*, cioè inerente allo svolgimento delle cose. L'utente del linguaggio non può operare performativamente sul significato delle parole, ma può solo scoprirlo e prenderne atto. In questa prospettiva, può esistere solo un unico e vero significato per ogni termine, che rispecchia la sua essenza ed è, quindi, predeterminato dalla realtà.

Le due concezioni del linguaggio, brevemente richiamate, influenzano in modo diverso l'uso della definizione. Come ha chiarito Herbert Hart «la definizione [...] consiste essenzialmente nel tracciare delle linee e nel distinguere tra un genere di cose e un altro, che il linguaggio designa con termini separati»⁷².

Se si segue la prospettiva convenzionalistica, la definizione, detta definizione nominalistica, verte su parole e se ne distinguono tre tipi: quella lessicale, quella stipulativa e quella esplicativa o ridefinitoria⁷³.

La definizione lessicale descrive l'uso di un termine in un dato gruppo di parlanti. La definizione stipulativa crea nuovi termini o modifica radicalmente il significato di termini esistenti tramite la stipulazione di un determinato significato. La definizione esplicativa o ridefinizione consente di delimitare o precisare il significato di un termine, in maniera funzionale agli obiettivi dell'universo discorsivo entro cui il termine è usato, mantenendo con l'uso ordinario un certo legame. Per l'approccio convenzionalistico al linguaggio, la definizione diviene uno strumento per fare un po' di ordine attraverso il controllo sui significati. Diventa un mezzo per superare alcuni problemi, già a suo tempo evidenziati da John Austin, ovvero: 1) se le parole sono strumenti dobbiamo cercare di usare *clean tools* per non cadere nelle trappole linguistiche; 2) se le parole non sono fatti o cose, dobbiamo essere in grado di astrarle dal mondo per individuare la potenziale inadeguatezza e arbitrarietà e rivolgerci al reale senza paraocchi⁷⁴. Lo

⁷² H.L.A. HART, *Il concetto di diritto*, Torino, Einaudi, p. 18.

⁷³ U. SCARPELLI, *La definizione nel diritto*, in ID. (a cura di), *Diritto e Analisi del Linguaggio*, Milano, Edizioni Comunità, 1976, pp. 183-197.

⁷⁴ J.L. AUSTIN, *A Plea of Excuses*, in *Proceedings of the Aristotelian Society*, 1956-1957, 57, pp. 1-30, ivi, p. 7 ss. Per un approfondimento sul ruolo di Austin nella filosofia linguistica, cfr. U. SCAR-

strumento definitorio consente di fare questa operazione, in quanto attraverso la delimitazione del significato possiamo sperare che dispute apparentemente irrisolvibili, derivanti da incomprensioni dovute all'uso del linguaggio, trovino la strada per la condivisione di soluzioni pratiche⁷⁵.

Muta completamente il ruolo della definizione se essa è correlata alla concezione essenzialistica del linguaggio. Ne muta l'oggetto, la natura, la struttura e la funzione. In questa prospettiva, la definizione, detta reale, verte su cose, può essere diretta o per genere prossimo e differenza specifica, si risolve in un discorso vero o falso ed ha funzione conoscitiva, perché consente di descrivere l'essenza delle cose di cui le parole sono lo specchio. In virtù della predeterminazione dei significati dei termini non sono possibili operazioni volte a cambiare o ridefinire il significato delle parole da parte dei parlanti. La definizione reale e la correlata concezione essenzialistica del linguaggio sono alla base delle ipostatizzazioni più perniciose che si contano oggi nella dialettica degli opposti. Sebbene sia la concezione essenzialistica sia la teoria della definizione reale siano state sottoposte a incisive critiche, rimangono subliminalmente diffuse, anche tra categorie di professionisti⁷⁶, e quindi a maggior ragione saranno radicate nella mente dell'uomo comune.

Così come le funzioni del linguaggio non si svolgono sempre in modo autonomo e ci sono situazioni in cui è difficile, se non si dispone di strumenti di analisi linguistica adeguati, separare il piano descrittivo da quello prescrittivo perché si insinua tra i due la valenza persuasiva, allo stesso modo la definizione dei termini può assumere i toni della persuasione. Attraverso la definizione persuasiva si cerca di avallare come certezze, quelle che in realtà sono credenze su basi soggettive, funzionali a interessi non palesemente dichiarati. Gli argomenti persuasivi possono, infatti, contare sulla definizione persuasiva di alcuni termini, il cui significato non è univoco ed ha assunto nel tempo una valenza emotiva o estremamente positiva o al contrario negativa. Nella storia dei termini che hanno subito questa sorte possiamo richiamare, a titolo esemplificativo, quello di eugenetica. Il suo significato lessicale si riferisce al ramo della genetica volta al progressivo miglioramento della specie umana attraverso l'incrocio di individui portatori di caratteri geneticamente favorevoli. Tuttavia, a seguito delle discriminazioni e degli abusi perpetrati sulla base di una mal interpretata politica eugenetica, il termine si è dotato di un significato persuasivo negativo da cui non si è più liberato.

PELLI, *John Langshaw Austin e il linguaggio comune*, in ID., *L'etica senza verità*, Bologna, Il Mulino, 1982, pp. 39-47.

⁷⁵ P. BORSELLINO, *Libertà, giustificazione della pena e metodo delle discipline penali in Calderoni*, in *Rivista Critica di Storia della Filosofia*, 1979, 3, pp. 313-349, ripubblicato in S. SALARDI, M. SAPORITI (a cura di), *Le tecnologie 'moralì' emergenti e le sfide etico-giuridiche delle nuove soggettività*, Torino, Giappichelli, 2020, pp. 90-124.

⁷⁶ M. JORI, A. PINTORE, *Manuale di teoria generale del diritto*, Torino, Giappichelli, 1995, p. 2. Gli autori hanno osservato che questa concezione «ha tuttora una certa diffusione anche presso i giuristi».

Le valenze persuasive unitamente alla persistente concezione essenzialistica del linguaggio, che fa coincidere il significato di una parola con una pretesa verità⁷⁷, inducendo a ritenere che concetti e categorie (anche giuridiche) siano «entità fisse, date una volta per tutte e indiscutibili, di cui possa essere fatta applicazione in tutti i campi»⁷⁸, sono parte costitutiva anche della narrazione dell'IA e fanno il gioco di chi ha interessi a imporre una strada, ossia quella della c.d. *governamentalità algoritmica*⁷⁹, senza dare alternative.

Ecco perché risulta di grande rilevanza anche pratica⁸⁰ la determinazione più chiara possibile del significato dei termini usati nel narrare l'IA nei diversi contesti discorsivi⁸¹.

La chiarezza passa, quindi, innanzitutto per l'abbandono definitivo della concezione essenzialistica del linguaggio, che purtroppo è ancora spesso sottesa non solo al dibattito comune sull'IA, ma altresì – e ben più grave – al dibattito giuridico.

Se si guarda al processo di individuazione dei significati dei termini con lo sguardo offuscato dalla concezione essenzialistica, si tenderà a lasciare in ombra la componente volitiva e decisoria che sta dietro l'uso di certi termini, poiché si sottolineerà un c.d. significato intrinseco o proprio del termine definito. Questo modo di procedere ha ricadute pratiche enormi. Consente innanzitutto a chi detiene non solo il potere tecnologico, ma altresì economico e con esso quello comunicativo, di perseverare nella rappresentazione antropomorfizzata dell'IA, che è funzionale a costruire un'illusoria fiducia nelle tecnologie, invece di mettere in primo piano il problema della fiducia nei confronti di chi le progetta, produce e commercia. Inoltre, consentirà agli attori del potere alcune operazioni, sottili e perniciose, quali: 1) mistificare il percorso del progresso tecnologico attraverso la retorica dell'inevitabilità; 2) oscurare le scelte etiche sottese alle scelte tecniche ed economiche; 3) deresponsabilizzarsi rispetto al ruolo di comunicatori di corretta e trasparente informazione circa la portata delle innovazioni, che mettono a disposizione degli individui e della macchina sociale.

In sintesi, molto di quello che, in tema di sviluppo tecnologico, può accadere o è già accaduto di pernicioso per la tenuta delle società democratiche e per il rispetto dei valori di eguaglianza e non discriminazione, si è verificato attraverso l'uso del linguaggio. In particolare, attraverso scelte che hanno privilegiato l'ambiguità e la vaghezza semantica, anziché un rigoroso controllo linguistico sul piano comunicativo.

È innegabile che non è facile governare il linguaggio in una società caratterizzata dalla velocità esponenziale della comunicazione. La velocità unitamente

⁷⁷ K. POPPER, *Scienza e Filosofia*, Torino, Einaudi, 1969, p. 102.

⁷⁸ P. BORSELLINO, *Bioetica tra 'moralì' e diritto*, Milano, Raffaello Cortina, 2018, p. 82.

⁷⁹ A. ROUVROY, B. STIEGLER, *Il regime di verità digitale*, cit.

⁸⁰ P. BORSELLINO, *Libertà, giustificazione della pena e metodo delle discipline penali in Calderoni*, cit.

⁸¹ P. BORSELLINO, *Bioetica tra 'moralì' e diritto*, cit. p. 483.

alla proliferazione di canali informativi, che affiancano quelli istituzionali, non solo non consente di porre sempre l'attenzione necessaria ad un'informazione linguisticamente accurata da parte di chi deve informare, ma inibisce altresì la richiesta di accuratezza da parte degli stessi utenti, che preferiscono *flash news* a presentazioni articolate. Il problema non è soltanto l'accuratezza grammaticale e sintattica. Il problema è quello dell'illusione informativa, poiché il destinatario è considerato più un interlocutore che va catturato, piuttosto che informato. In questo scenario, la valenza persuasiva gioca un ruolo di primo piano.

La sua presenza silenziosa inficia la comunicazione di informazioni su tematiche complesse e dal forte impatto sulla vita del destinatario dell'informazione. Comunicare questioni tecniche al pubblico non significa usare un linguaggio difficile, ma evitare un linguaggio finto facile che fa uso di parole semplici o meglio semplificate, ma che a causa dell'eccesso di uso hanno perso il loro potere informativo. Per permettere al destinatario dell'informazione di riuscire a formulare un suo pensiero su questioni che incidono sulla sua vita, ma che sono caratterizzate da un alto grado di tecnicismo, non è al "parlare facile" nel senso banalizzante che bisogna fare riferimento. Bisogna perseguire la via della chiarezza che illumina, che non è ingannevole né compiacente, né sbrigativa o semplificante.

Quello proposto è certamente un obiettivo ambizioso, ma che, tutto considerato, vale la pena di provare a perseguire, nei limiti del possibile, se vogliamo essere individui che fanno la loro parte nel momento storico in cui vivono e con ciò incidere sulle condizioni in cui vivranno le generazioni future. Per perseguire la trasparenza e la chiarezza discorsiva occorre riflettere sui concetti usati nella narrazione dell'IA, partendo dall'idea che chi la qualifica come rivoluzione sta facendo un'operazione persuasiva e si avvale, pertanto, della vaghezza e dell'ambiguità di molti altri termini, che hanno accompagnato lo sviluppo delle tecnologie intelligenti. Possiamo, pertanto, parlare di tassonomia concettuale della visione antropomorfa dell'IA.

4. *Tassonomia concettuale dell'antropomorfizzazione dell'IA*

Il tema dell'antropomorfismo del macchinico è, già da qualche decennio, affrontato criticamente in letteratura. Frank Pasquale, ad esempio, sottolinea il grande pericolo di un antropomorfismo che definisce ingannevole, in quanto si realizza per il tramite dell'educazione ed esposizione dei bambini a insegnanti robot a scopo di «un indottrinamento subdolo per insegnare ai bambini che umano e macchina sono in fondo uguali e interscambiabili»⁸². L'antropomor-

⁸² F. PASQUALE, *Le nuove leggi della robotica. Difendere la competenza umana nell'era dell'intelligenza artificiale*, Roma, LUISS University Press, 2021, p. 90. L'autore specifica a pagina 280 alla nota 32 che secondo la sua opinione «la chiave del classico "test" di Turing sia educare ad

fizzazione delle macchine presenta, infatti, molti volti e può perpetuarsi attraverso diversi canali. Tuttavia, il nodo cruciale di questo fenomeno, che gli consente di essere ubiquo a livello discorsivo, è dato dall'impiego di specifici concetti nella narrazione della storia dell'IA. Di fatto, la tassonomia concettuale di questa storia alimenta la dialettica degli opposti e rende arduo ogni tentativo di fare ordine in questa narrazione. Per economia discorsiva dovremo tagliare attraverso la fitta rete di termini antropomorfizzanti per andare alla radice del problema. Nella scelta di quali termini analizzare sul piano semantico si è deciso di valutare le seguenti caratteristiche: 1) la possibilità di essere usati in modo ubiquo in diversi universi discorsivi, dal linguaggio ordinario ai linguaggi specialistici; 2) la caratterizzazione dei termini come normativi e l'essere privi di un referente semantico nella realtà; 3) il fatto di avere acquisito con l'uso una forte valenza persuasiva; 4) l'essere riferiti storicamente all'uomo per identificare sue qualità distintive rispetto ad altri enti, soprattutto in relazione alle caratteristiche del suo agire (libero arbitrio). I termini in possesso di queste caratteristiche, che ci apprestiamo ad analizzare, sono quelli di autonomia, persona, responsabilità, imprevedibilità e intelligenza. Essi formano le basi del processo di antropomorfizzazione del macchinico e sostengono la visione dicotomica uomo-macchina. Sono tutti termini con una risalente storia semantica e quasi tutti accomunati dall'essere protagonisti del tradizionale dibattito attorno alla libertà di agire (il dibattito sul libero arbitrio). Solo il termine intelligenza non è stato esplicitamente parte attiva della risalente disputa tra sostenitori del libero arbitrio e coloro che, invece, promuovono una visione deterministico-riduzionistica dell'agire umano. Tale termine si inserisce però prepotentemente nell'attuale dialettica degli opposti. Anzi, forse ne ha permesso la larga diffusione e il radicamento. Di fatto, tutta la storia dell'IA, da quando è stato coniato il termine intelligenza artificiale nel 1956 in occasione del *Dartmouth Summer Project on Artificial Intelligence*, che rappresentò il momento inaugurale della cibernetica, ruota attorno alla definizione dell'IA *per relationem* rispetto all'intelligenza umana. Se pensiamo, ad esempio, a uno dei grandi protagonisti dello sviluppo di queste macchine, Alan Turing, ci si avvede come già nel suo pensiero sia radicata questa tendenza. Nel suo articolo *Computing Machinery and Intelligence* del 1950 affronta il dilemma relativo alla creazione di macchine pensanti, quindi dotate di una caratteristica umana, e non esclude che ciò possa realizzarsi⁸³.

La nozione di intelligenza si rivela di cruciale importanza nella narrazione

accettare le AI come persone: non si tratta di un sistema filosofico per determinare se una AI è ben fatta, ma di un artificio retorico per abituare i lettori all'idea che sapere svolgere determinate azioni permetta di essere riconosciuti come persone».

⁸³ Cfr. A.M. TURING, *Macchine calcolatrici e Intelligenza*, in V. SOMENZI, R. CORDESCHI (a cura di), *La filosofia degli automi. Origini dell'intelligenza artificiale*, Torino, Paolo Boringhieri, 1986, pp. 157-183, ivi, p. 173.

dell'IA. Essa fa parte della griglia di concetti *primi*, per mezzo dei quali ha potuto prosperare la dialettica degli opposti e l'antropomorfismo della tecnica.

I concetti *primi* sono nozioni che designano qualità delle azioni e/o dell'agente. Si dice, ad esempio, "ha agito autonomamente, responsabilmente, imprevedibilmente, intelligentemente" e anche "Tizio è una persona responsabile, imprevedibile, autonoma, intelligente".

Sono termini con una storia semantica antica, motivo per cui il loro impiego richiede un chiarimento sulle condizioni d'uso, ossia sulle classi di cose o fatti denominati dalla parola (estensione) e sull'insieme delle proprietà che devono accomunare certe cose o fatti affinché appartengano alla classe denotata (intensione). Queste operazioni definitorie, che proporremo per i termini richiamati, sono fondamentali per comprendere la dialettica degli opposti.

4.1. *Intelligenza, autonomia e imprevedibilità della persona umana*

Consideriamo dapprima i termini citati quando sono usati con riferimento all'essere umano e cominciamo con quello di intelligenza. La parola intelligenza nel sentire comune si riferisce al possesso di diverse proprietà che vanno da quelle più strettamente razionali a quelle che fanno leva sugli aspetti emotivi. L'intelligenza è definita spesso da aggettivi qualificatori come in espressioni "avere un'intelligenza pratica", "un'intelligenza speculativa", o ancora "rara intelligenza", "intelligenza acuta, fiacca, tarda" ecc. L'ambito di facoltà mentali e psichiche attribuite all'uomo, che riassumiamo con questo termine, è quindi molto ampio. Se vogliamo circoscrivere le condizioni minime del suo uso possiamo richiamare le capacità psichiche e mentali, che consentono di ragionare, comprendere la realtà e inserirsi in un contesto e ivi orientarsi. Tuttavia, l'intelligenza umana non si limita alle facoltà mentali così intese, ma include la corporeità e l'emotività. La ricchezza di sfumature semantiche della parola intelligenza unitamente alla sua valenza persuasiva positiva può facilmente indurre fraintendimenti se chi adopera il termine non ridefinisce le condizioni d'uso in un dato contesto.

Quando fu coniata l'espressione *Artificial Intelligence* forse non ci si è soffermati a riflettere sulle implicazioni di questa scelta linguistica e forse nemmeno si immaginava che avrebbe alimentato una dialettica degli opposti e un profondo approccio dicotomico tra uomo e macchinico⁸⁴. La scelta del nome da dare nel 1956 alle nuove macchine ha però innescato quell'ideologia antropomorfizzante, che oggi può venire sfruttata da chi vuol fare apparire le macchine migliori degli uomini (lo chiamerò "antropomorfismo in eccedenza").

⁸⁴ Tale dialettica si avvale anche di metafore belliche, basti pensare al titolo del volume di L. ALEXANDRE, *La guerra delle intelligenze. Intelligenza artificiale contro intelligenza umana*, cit.

Se la nozione dell'intelligenza è il concetto primo della narrazione dell'IA, allora merita attenzione istituzionale e un'adeguata ridefinizione. Ed in effetti questo è quello che è accaduto con il lavoro intitolato *A Definition of AI: Main Capabilities and Disciplines* del 2019, elaborato dall'*High-Level Expert Group on Artificial Intelligence*. Il documento si propone di elaborare una definizione di IA che chiarisca certi aspetti della disciplina in oggetto, il fine è quello di «avoid misunderstandings, to achieve a shared common knowledge of AI that can be fruitfully used also by non-AI experts, and to provide useful details that can be used in the discussion on both the AI ethics guidelines and the AI policies recommendations»⁸⁵.

Il Gruppo di esperti parte proprio da considerazioni legate alla vaghezza del termine intelligenza e ne circoscrive il significato. Si riconosce che i contenuti dell'intelligenza sono stati oggetto di studio in varie discipline e ci si concentra sul fatto che, nel campo delle ricerche sull'IA, il termine si riferisce alla razionalità. La definizione proposta dal Gruppo, infatti, parte da quella elaborata dalla Commissione europea, secondo la quale «AI refers to systems that display intelligent behaviour by analysing their environment and taking actions – with some degree of autonomy – to achieve specific goals»⁸⁶. Il Gruppo precisa ulteriormente questa definizione intendendo per sistema di IA «any AI-based component, software and/or hardware. Indeed, usually AI systems are embedded as components of larger systems, rather than stand-alone systems»⁸⁷. Il Gruppo sottolinea come l'IA sia un sistema razionale e spiega come raggiunge questa razionalità «perceiving the environment in which the system is immersed through some sensors, thus collecting and interpreting data, reasoning on what is perceived or processing the information derived from this data, deciding what the best action is, and then acting accordingly, through some actuators, thus possibly modifying the environment. AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behaviour by analysing how the environment is affected by their previous actions»⁸⁸. Il Gruppo propone, infine, una definizione aggiornata di IA come «Artificial intelligence (AI) systems are software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the in-

⁸⁵ HIGH-LEVEL EXPERT GROUP ON ARTIFICIAL INTELLIGENCE, *A Definition of AI: Main Capabilities and Disciplines*, Brussels, European Commission, 2019, p. 1, accessibile al link <https://digital-strategy.ec.europa.eu/en/library/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines>, ultimo accesso agosto 2022.

⁸⁶ *Ibidem*, p. 3.

⁸⁷ *Ibidem*.

⁸⁸ *Ibidem*.

formation, derived from this data and deciding the best action(s) to take to achieve the given goal. AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behaviour by analysing how the environment is affected by their previous actions»⁸⁹.

Questa ridefinizione istituzionale rappresenta un importante passo verso la ricerca di chiarezza e trasparenza attorno alle potenzialità dell'IA. Da tale ridefinizione emerge che l'IA è un sistema, prodotto dall'uomo, in grado di portare a termine un compito complesso, utilizzando sensori con cui indaga l'ambiente e attuatori con cui agisce su di esso. Resta fuori, da questa definizione, il riferimento all'intelligenza emotiva, mentre si punta al collegamento con il significato di *razionalità*. In tal modo si mira a esaltare quel senso di intelligenza come razionalità, per il quale rilevante è la commisurazione dell'agire «in base all'efficacia dei mezzi prescelti in vista di un fine»⁹⁰. È, quindi, una ridefinizione che mette in luce uno degli aspetti dell'intelligenza che possono avere sia uomini che macchine, e indirettamente consente di rivalutare l'intelligenza umana alla luce di ulteriori criteri che la contraddistinguono. Va, infatti, considerato che, sebbene nel linguaggio comune il concetto di razionalità sia usato come sinonimo di ragione e l'aggettivo razionale come sinonimo di ragionevole, in realtà i due concetti hanno sfumature molto diverse sul piano filosofico. La nozione di ragione, pur avendo più significati, ha un nucleo di significato che la definisce come «guida autonoma dell'uomo in *tutti* i campi nei quali un'indagine o una ricerca è possibile» e, in quanto tale rappresenta «la forza che libera dai pregiudizi, dal mito, dalle opinioni radicate ma false, dalle apparenze e consente di stabilire un criterio *universale* o *comune*, per la condotta dell'uomo in tutti i campi»⁹¹. Con la ridefinizione di intelligenza come razionalità, ma non come ragione, si circoscrive molto l'ambito di applicazione del termine⁹². Ciò a dire che dell'IA si può predicare l'intelligenza come razionalità, ma non come ragione. L'IA può considerarsi dotata di una razionalità ben precisa, legata ad azioni determinate dai dati raccolti e interpretati secondo registri interpretativi prestabiliti, con un più o meno ampio margine di operatività, che viene comunque deciso a monte, da chi programma il sistema. Non sarà, almeno allo stato attuale dello sviluppo tecnologico, una intelligenza anche *ragionevole*, ossia basata sulla ragione secondo la definizione poc'anzi richiamata. Il chiarimento di questo punto offre uno strumento di contrasto di quelle visioni dell'IA come intelligente in

⁸⁹ *Ibidem* p. 8.

⁹⁰ N. ABBAGNANO, *Dizionario di Filosofia*, cit. p. 900.

⁹¹ *Ibidem*, p. 892.

⁹² Alcuni autori hanno sottolineato che non sarebbe necessario il requisito della razionalità nella definizione di IA e che, comunque, questa va intesa come un *outcome* e non come un processo, cfr. A. RENDA, *Moral Machines. The emerging EU policy on "Trustworthy AI"*, in W. BARFIELD (eds.), *The Law of Algorithms*, Cambridge, Cambridge University Press, 2021, pp. 667-690, ivi, p. 669.

eccedenza (antropomorfismo in eccedenza) e consente di rivalutare l'intelligenza umana, ricollocandola al centro dell'attenzione in quanto meritevole, non tanto di essere saccheggiata per dare alle macchine qualcosa di simile, ma di essere nutrita, non solo per fornire migliori prestazioni intellettuali, ma per garantire il pluralismo delle sue caratteristiche e con esso l'uguaglianza tra esseri umani, segnando la differenza in eccedenza a favore dell'uomo e non del macchinico.

L'essere umano è, quindi, dotato di ragione, da intendersi come forza che è alla base del giudizio e del discernimento. Con la ragione, pienamente sviluppata nel senso kantiano, l'uomo può ottenere autonomia e libertà.

Siamo così giunti a uno dei più controversi concetti impiegati nella dialettica degli opposti, ossia quello di autonomia. Il termine compare in particolare quando si discute delle auto a guida autonoma⁹³ o di robot autonomi, capaci di autoapprendimento e in quanto tali imprevedibili nelle loro azioni. Il concetto di autonomia ha una lunga storia nel pensiero filosofico ed è, da qualche decennio, al centro del dibattito su questioni eticamente rilevanti quali, ad esempio, quelle che concernono il fine vita, l'ambito procreativo, la genetica e le neuroscienze. Nella storia semantica del concetto, il suo significato è applicato esclusivamente agli esseri umani⁹⁴ ed è strettamente connesso sia a quello di libertà *da* costrizioni e impedimenti esterni sia a quello di libertà *di*, ossia capacità e potere di compiere scelte. Questa distinzione è stata affrontata da Isaiah Berlin, il quale nel distinguere le due accezioni, chiarisce che "libertà *da*" (libertà negativa) delimita «l'area entro cui si permette o si dovrebbe permettere al soggetto di fare o essere ciò che è capace di essere e di fare, senza interferenze da parte di altre persone»⁹⁵. Nella seconda accezione di libertà *di* (libertà positiva), invece, l'autonomia che si esprime mediante tale collegamento alla libertà positiva è quella di essere padroni di sé stessi, potendo esprimere il potere di compiere scelte autonome.

Le due accezioni di autonomia richiamate includono, a loro volta, il riferimento alla nozione di responsabilità e a quello di imprevedibilità. Va da sé che se un soggetto è autonomo nelle scelte, egli dovrà farsi carico delle loro eventuali conseguenze sugli altri, sarà pertanto considerato responsabile. Una decisione umana autonoma è una decisione che è presa *all things considered*, ossia rispettando certe condizioni tra cui: assenza di costrizioni fisiche e psicologiche; presenza della capacità di comprensione di uno spettro di informazioni dispo-

⁹³ Sono termini che ricorrono nei discorsi tecnici, in quelli istituzionali e nel linguaggio ordinario, determinando, attraverso questa ubiquità, un forte rischio di fraintendimenti grossolani.

⁹⁴ Secondo Laura Palazzani «[a]utonomy in its ethical original meaning can be attributed only to human beings [...]. The term "autonomy" cannot be applied to artefacts, even if very advanced complex or even "intelligent" systems», L. PALAZZANI, *AI and health: ethical aspects for regulation*, in *Teoria e Critica della Regolazione Sociale*, 2021, 1, pp. 1-16, ivi, p. 4.

⁹⁵ I. BERLIN, *Due concetti di libertà*, in H. HARDY (a cura di), *Libertà*, Milano, Feltrinelli, 2005, p. 172.

nibili e rilevanti che rendono la scelta decisoria non manipolata; capacità di operare la scelta – a seconda della pregnanza della situazione – sulla base non solo di considerazioni utilitaristiche e di opportunità, ma integrando nel processo deliberativo il bilanciamento tra valori espressi da norme morali e/o giuridiche, che fungono da guida dei propri e altrui comportamenti⁹⁶. Quando si parla di *scelte di valore* si intende proprio questo. I valori a cui si ispirano le scelte possono mutare nel tempo e le scelte operate possono, quindi, essere riviste alla luce di nuove ponderazioni. Pertanto, l'autonomia si esprime per l'appunto nella possibilità di cambiare posizione, assumendosene la responsabilità. Questa responsabilità potrà essere morale, giuridica, politica e sociale. Il concetto di responsabilità chiama sempre in causa norme di condotta per la sua definizione. È un concetto normativo in cui sono sempre presenti elementi che possiamo definire artificiali, nel senso di determinati artificialmente dalla società. L'uso di questo concetto, come ha ben chiarito Uberto Scarpelli⁹⁷, richiede che siano sempre presenti tre elementi: un dovere, una conseguenza (sanzione) e un comportamento.

Chi è responsabile per le proprie azioni è un agente o persona morale. La sua responsabilità rispetto al dovere da rispettare può essere modulata a seconda che vi siano impedimenti patologici o situazioni che richiedono una valutazione della sua capacità di agire autonomamente. I sistemi giuridici moderni assumono la capacità piena in ogni agente o persona e prevedono eccezioni solo in casi particolari e da dimostrare. È quindi sulla portata della conseguenza che si lavora quando si prova che una persona non aveva la piena o totale capacità, non sulla modulazione del dovere previsto dalla norma di condotta, che resta invariato. Attraverso lo schema della responsabilità si instaura, quindi, un rapporto funzionale tra soggetto e azioni che realizza. In pratica, questo schema di responsabilità è un *costrutto sociale*, ossia «un legame stabilito socialmente in relazione anche a criteri culturali»⁹⁸. A tale schema è sottesa la ricerca di un ordine

⁹⁶ Cfr. N. BOBBIO, *Norma*, in ID., *Contributi ad un Dizionario Giuridico*, Torino, Giappichelli, 1994, pp. 177-214. L'autore, tra le altre cose, analizza la teoria di Hart del c.d. buon cittadino.

⁹⁷ U. SCARPELLI, *Riflessioni sulla responsabilità politica. Responsabilità, libertà, visione del mondo*, in R. ORECCHIA (a cura di), *La responsabilità politica, Diritto e Tempo*, Atti del XIII Congresso nazionale, Milano, Giuffrè, 1982, pp. 41-95. Scarpelli definisce la responsabilità nel seguente modo «[d]ico di un soggetto che ha od aveva responsabilità se ha, o aveva, un dovere di comportamento; se è a lui eventualmente riferibile, o attualmente riferito, un comportamento in sé stesso oppure in quanto produttivo di certi effetti, contrastante con il dovere, e pertanto oggetto di valutazione negativa; se, in dipendenza dal riferimento del comportamento oggetto della valutazione negativa, è a lui eventualmente imputabile o attualmente imputata, una conseguenza, a sua volta oggetto di una valutazione negativa», *ivi*, p. 47.

⁹⁸ Cfr. A. CERETTI (1990), *È possibile parlare di responsabilità morale in criminologia?*, in A. CERETTI, I. MERZAGORA (a cura di), *Criminologia e Responsabilità Morale*, Padova, Cedam, 1990, pp. 241-265, *ivi* p. 261.

sociale in cui si orientano le condotte individuali in modo da circoscrivere o evitare comportamenti imprevedibili.

L'imprevedibilità delle azioni umane può essere interpretata in due opposti sensi. In un primo senso, un soggetto imprevedibile può essere colui che è *fuori controllo*, ossia mette in atto comportamenti che violano le norme di condotta condivise. In un secondo senso, un soggetto può essere imprevedibile in quanto è capace di sorprendere con idee e comportamenti *fuori dal comune*, ma non *fuori controllo*. In questo secondo caso, lo schema regolativo delle condotte sociali viene rispettato, ma la condotta posta in essere non è quella che usualmente ci si sarebbe aspettati. Quando un soggetto è *fuori controllo* vanno indagate le ragioni e va verificata la possibilità che si possa in qualche modo assumere la responsabilità per la violazione dello schema regolativo socialmente apprestato. Nel caso in cui un soggetto agisca, invece, *fuori dal comune* non vanno indagate né le ragioni dell'agire né va verificato lo stato mentale perché risulta pienamente responsabile, essendo il suo comportamento coerente con le norme di condotta condivise.

La responsabilità morale e giuridica per le proprie azioni e, quindi, l'autonomia soggettiva è un riflesso della capacità di *agire diversamente*. Nella letteratura filosofica del secolo scorso, ci si è molto occupati della responsabilità e della libertà di agire, attraverso l'analisi di una proposizione precisa: "poteva avere agito diversamente". Nell'analizzare tale proposizione, Alf Ross ha puntualmente messo in evidenza che definire la possibilità o meno di agire diversamente dipende dalla co-presenza di tre condizioni. Sono necessarie: condizioni *costruttive o costituzionali*, ossia che non vi siano costrizioni dovute, ad esempio, a patologie mentali; condizioni *occasional*i, ossia vi devono essere le circostanze che consentono l'azione; condizioni *motivazionali*, ossia rappresentano la volontà o il motivo dell'azione⁹⁹. Secondo l'analisi di Ross, data la costituzione e l'occasione, quando diciamo che un soggetto "non avrebbe potuto agire diversamente", riferita a situazioni coattive, stiamo in realtà dicendo che "non potevamo pretendere che agisse diversamente, benché avrebbe potuto". Si pone, pertanto, l'accento sulla volontà e sulla capacità di decidere autonomamente.

Come osservato, dal punto di vista morale e giuridico, l'esercizio dell'autonomia e l'attribuzione di responsabilità presuppongono un agente morale o persona. Chiamano, pertanto, in causa la nozione di persona, che ha, come gli altri concetti *primi*, una lunga storia semantica durante la quale si sono stratificati diversi significati. Trattasi di un concetto normativo che richiama il trattamento normativo correlato all'ascrizione di diritti e doveri. Persona non è l'individuo in senso bio-psicologico, ma è il trattamento normativo o se si vuole l'insieme di norme che, in un dato momento storico, sono predisposte a favore o contro cer-

⁹⁹ A. ROSS, *Colpa, responsabilità e pena*, Milano, Giuffrè, 1972, pp. 255-285.

te linee di condotta¹⁰⁰. Ciò spiega come mai non sempre coloro che sul piano descrittivo appartengono al genere umano sono stati trattati come persone sul piano morale e giuridico. Si pensi alla schiavitù o al nascituro non è persona sul piano giuridico, ma soggetto di diritto. È stata la separazione tra diritto e natura e tra diritto e morale, frutto della riflessione giuspositivistica del secolo scorso, che ha contribuito in maniera decisiva a emancipare il concetto di persona da referenti fattuali. A ciò si è aggiunto un ulteriore importante passaggio, prodotto del costituzionalismo moderno (c.d. personalismo costituzionale), quello di collegare la persona all'eguaglianza non solo formale, ma anche sostanziale, superando in tal modo il «problema della compatibilità tra soggetto astratto e riconoscimento delle differenze»¹⁰¹. L'eguaglianza nei diritti consente di attribuire uguale valore alle differenze e «forma il “valore” o la “dignità” della persona»¹⁰². All'attuale nozione di persona e all'uso che se ne fa nel contesto giuridico attuale sono, quindi, sottese precise scelte di valore. Con tale termine si prende partito per una visione inclusiva e antidiscriminatoria della società. Non è necessariamente una nozione che deve essere preconceputamente chiusa all'inclusione di altri enti, come ad esempio, la natura o l'IA. Quello che il significato etico e giuridico attuale del termine persona ci segnala è che la sua estensione comporta anche l'estensione della dimensione egualitaria e va pertanto ben ponderata, quando si vogliono includere altri enti.

Da quanto precede emerge che le nozioni analizzate in questo paragrafo sono storicamente riferite all'essere umano, non perché vi sia una qualche definizione vera, ma perché, indipendentemente dalle sfumature di significato assunte con il trascorrere del tempo e l'uso protratto, queste parole si sono trovate a raccontare l'evoluzione concettuale dei comportamenti umani in società.

A muovere da tali premesse, occorre chiarire come queste stesse nozioni vengano impiegate nella narrazione attorno all'IA e in che modo alimentino la dialettica degli opposti e la visione antropomorfa della tecnica¹⁰³.

¹⁰⁰ H. KELSEN, *Lineamenti di dottrina pura del diritto*, cit.

¹⁰¹ S. RODOTÁ, *Il diritto di avere diritti*, cit. p. 147.

¹⁰² L. FERRAJOLI, *Principia iuris. Teoria del diritto e della democrazia*, vol. II, Roma-Napoli, Laterza, 2007, p. 587.

¹⁰³ In tal senso, il linguaggio mediatico impiega tantissimi termini, che si riferiscono al mondo umano, per parlare dell'IA. Tale linguaggio non si limita all'impiego dei concetti *primi* di cui abbiamo dato conto. Cfr. i titoli dei seguenti quotidiani: 1° maggio 2021, *Il Giorno* di Milano, *L'algoritmo ci ha intrappolato fuori provincia* e ancora *Noi prof. intrappolati dall'algoritmo impazzito*; 14 settembre 2018, *Il Venerdì* di Repubblica, *La parola ai robot*; il 20 giugno 2018, *Corriere della Sera*, *Uomo contro macchina. La prima sfida di dialettica finisce con un pareggio*; *Panorama* del 15 luglio 2018, *E se l'Europa attribuisse una “personalità elettronica” ai robot?*; 12 luglio 2017, *Daily Mail*, *Rise of Robots could cost 15 million UK jobs and will ‘break the social ladder’ as automation hits poorer workers hardest, report finds*.

4.2. *Intelligenza, autonomia e imprevedibilità della persona artificiale*

Con riguardo alla nozione di intelligenza, abbiamo dato conto dell'intervento definitorio delle istituzioni europee, avente lo scopo di chiarire e rendere comprensibili, anche ai non esperti, le realistiche prospettive di applicazione di ciò che definiamo macchine intelligenti. A fronte di questo importante sforzo istituzionale, che avrà auspicabilmente ricadute virtuose sulla dialettica degli opposti che pervade la narrazione mediatica dell'IA, quello che occorre fare ora è spingere affinché questo controllo linguistico trovi sistematica implementazione per correggere la rotta verso un dibattito e un racconto dell'IA privo di grossolani fraintendimenti.

La proposta di Kirchsclaeger¹⁰⁴ di sostituire l'uso del termine IA con quello di sistemi basati sui dati (*data-based system*) va in questa direzione ed è condivisibile nella misura in cui risulta funzionale a «lifts the veil of the inappropriate attribution of the myth of “intelligence” covering substantial problems and challenges of data-based systems and allows for more accurateness, adquacy, and precision in the critical reflection of data-based systems»¹⁰⁵. L'autore ritiene che questa nozione sia più adatta a spiegare che cosa effettivamente costituisca l'IA, perché l'espressione scelta punta a sottolineare che si tratta di processi di generazione, raccolta, e valutazione dei dati, da cui dipendono la percezione basata sui dati, le predizioni basate sui dati, fino alle decisioni basate sui dati. L'espressione *data-based systems*, in poche parole, richiama direttamente ed esplicitamente il nodo centrale dell'IA, che è rappresentato dal poter disporre e utilizzare una enorme quantità di dati. Se per questa nozione vi sono, quindi, già proposte istituzionali e dottrinali alternative al suo impiego, lo stesso non si può dire per gli altri concetti *primi* impiegati nella narrazione. Per preparare il cammino verso una loro ridefinizione, proponiamo di seguito alcune riflessioni anche sui concetti di autonomia, imprevedibilità, persona e responsabilità quando sono impiegati con diretto riferimento alle tecnologie intelligenti.

Partiamo dalla nozione di autonomia. Nel paragrafo che precede, abbiamo visto cosa è incluso nel suo pendolo semantico quando il discorso verte sull'autonomia dell'essere umano. Qui ci chiediamo cosa si intende dire quando si usano le espressioni robot autonomi, auto autonome e IA autonoma.

L'uso del termine autonomia in relazione alle tecnologie intelligenti richiama il senso di autonomia sia come potere di agire senza impedimenti esterni nello spazio fisico in cui l'IA si trova ad operare sia come potere di scelta di quell'azione che, secondo un calcolo di vantaggio/svantaggio, rischio/benefi-

Anche Frank Pasquale denuncia che «[l']entusiasmo dei manager ha trascinato fin troppi giornalisti, che hanno scritto di “avvocati robot” e “medici robot” come se fossero già una realtà», F. PASQUALE, *Le nuove leggi della robotica*, cit. p. 32.

¹⁰⁴ P.G. KIRSCHSCHLAEGER, *Digital Transformation and Ethics*, Baden-Baden, Nomos, 2021.

¹⁰⁵ *Ibidem*, p. 103.

cio, meglio consente di raggiungere l'obiettivo pre-programmato dall'uomo. Possiamo dire che il punto di discriminazione tra l'azione autonoma umana e quella meccanica sta proprio nel modo in cui intendiamo il potere di scelta: la scelta dell'IA è una scelta in senso debole. Se è vero, infatti, che il calcolo rischio-beneficio è fatto dall'IA, l'obiettivo e lo scopo per cui il calcolo è effettuato sono però eteroimposti dall'uomo. Al contrario, la scelta umana è una scelta in senso forte, basata cioè sulla deliberazione effettuata in base alla ponderazione di valori introiettati a cui si ispira il giudizio sul proprio e l'altrui comportamento. Se, come nel caso dell'IA, l'obiettivo è eteroimposto, altrettanto lo sono gli eventuali valori sottesi al suo raggiungimento. Nel caso, ad esempio, delle *autonomous driverless cars*, la previsione delle modalità di decisione in situazioni caricate di un gravoso fardello morale, come è il caso del noto dilemma del carrello¹⁰⁶, sono predefinite dal programmatore e l'IA potrà scegliere solo quelle opzioni, a cui sono sottese scelte morali, che le sono state messe a disposizione¹⁰⁷. Diversamente dalle decisioni umane definite autonome in senso forte, resta fuori da quelle del meccanico la deliberazione effettuata sulla base di valori non pre-programmati. Se non si tiene conto di questa fondamentale differenza, il rischio è quello di far passare il messaggio che l'autonomia umana e quella delle macchine sono la stessa cosa, sono equiparabili. Questa operazione è perniciosa perché si risolve sempre a sfavore della valorizzazione del complesso mondo morale che caratterizza l'agire umano. Quest'ultimo, infatti, è un universo in cui si mettono in atto complessi processi deliberativi, frutto di un lungo ed elaborato sviluppo della coscienza morale individuale. In questo processo deliberativo sono incorporati criteri diversi che concorrono alla decisione finale, in forme e modi ancora non del tutto conosciuti¹⁰⁸. Vi contribuiscono emozioni e sentimenti, ma altresì credenze e stati conativi, come già sostenuto a suo tempo da filosofi come Mario Calderoni. Allo stato attuale delle conoscenze scientifiche sulla formazione dei giudizi morali, appaiono ancora non smentite le intuizioni di Calderoni sull'architettura delle azioni moralmente rilevanti. A tale architettura contribuirebbero congiuntamente sia le emozioni, che Calderoni denomina azioni riflesse e istintive,

¹⁰⁶ Formulato da Philippa Ruth Foot nel 1967.

¹⁰⁷ S. SALARDI, *Autonomia e responsabilità alla prova di autonomous artificial systems: Profili filosofici, etici e giuridici. Il caso delle driverless cars*, in D. CERINI, A. PISANI TEDESCO (a cura di), *Smart mobility, smart cars e intelligenza artificiale: responsabilità e prospettive*, Torino, Giappichelli, 2019, pp. 117-139.

¹⁰⁸ Le neuroscienze sono ancora agli inizi degli studi di questi processi e diverse sono altresì le posizioni filosofiche in materia. Allo stato attuale degli sviluppi nell'ambito delle scienze cognitive non è ancora possibile definire il *locus* del libero arbitrio. A fronte dello stato dell'arte scientifico appare del tutto inopportuno trarre conclusioni sul piano normativo che muovano verso visioni riduzionistiche e deterministiche dell'agire umano. Cfr. S. SONGHORIAN, *Etica e scienze cognitive*, Roma, Carocci, 2020.

sia il ragionamento, che per l'autore è la volontà¹⁰⁹. Questa posizione è condivisa oggi da filosofi come Shaun Nichols e Hanno Sauer, che, sebbene con importanti diversità nelle loro proposte, quello che condividono è «il fatto di proporre un modello» in materia di giudizi morali «che, da un lato, tiene conto dei risultati dei lavori sperimentali e li mette in dialogo con la tradizione filosofica, dall'altro rende conto sia del contributo della ragione sia di quello di sentimenti ed emozioni»¹¹⁰. A muovere da queste premesse si assume ancora oggi che, escludendo situazioni patologiche, l'essere umano possa agire diversamente e che lo possa fare sia perché vi è una componente di imprevedibilità *creativa* nel suo agire sia perché è *free to fall*¹¹¹.

Mentre quest'ultima componente dell'agire è preclusa all'IA, a meno che non la si voglia programmare per fare *del male*, resta da chiarire il ruolo e il significato della nozione di imprevedibilità quando riferita al macchinico. Quando si sostiene che l'IA è imprevedibile, si sta indicando una imprevedibilità tecnica¹¹², non riconducibile a scelte deliberative autodeterminate nei sensi appena richiamati. L'imprevedibilità di queste macchine è all'origine eterodiretta dalle scelte di programmare i sistemi in un certo modo e di immettere certi dati valoriali e di azione e di scartarne altri¹¹³.

Puntualizzare queste cruciali differenze semantiche è un modo per scardinare la dialettica degli opposti non solo a livello comunicativo e mediatico, bensì anche nei contesti specialistici in cui purtroppo si è insinuata. L'ubiquità a cui sono esposti i concetti *primi* sinora analizzati ha agevolato il dilagare della dialettica degli opposti che sostiene l'ideologia antropomorfizzante delle macchine. Vi sono contesti che, più o meno consapevolmente, per le loro caratteristiche, rischiano addirittura di perpetuare, nobilitandolo, l'antropomorfismo tossico. È questo il caso del diritto, in cui si è, da tempi risalenti, giocata la partita tra visioni riduzionistiche dell'umano e quelle volte alla valorizzazione della libertà. Di questo però si parlerà nel capitolo dedicato alle questioni strettamente giuridiche.

¹⁰⁹ M. CALDERONI, *I postulati*, cit. p. 103. L'autore affronta il tema della contrapposizione della volontà alle emozioni nel configurare forme di responsabilità penale sostenendo che «[...] le azioni riflesse ed istintive, che alcuni sogliono contrapporre alla volontà come qualche cosa di irriducibilmente diverso, siano invece il *materiale* onde la volontà si vale e senza il quale sarebbe impossibile comprendere i suoi movimenti».

¹¹⁰ S. SONGHORIAN, *Etica e scienze cognitive*, cit. p. 105.

¹¹¹ Cfr. J. HARRIS, *How to be good. The Possibility of Moral Enhancement*, Oxford, Oxford University Press, 2016.

¹¹² Cfr. S. SALARDI, *La dimensione semantica delle tecnologie intelligenti*, in *Notizie di Politeia*, 2021, 143, pp. 156-161.

¹¹³ Spunti di riflessione sul tema sono presenti anche in S. SALARDI, *L'attualità del pensiero di Mario Calderoni: questioni bioetiche di frontiera come case study*, in L. FORNI, S. SALARDI, M. SAPORITI (a cura di), *Scritti in onore di Patrizia Borsellino*, Torino, Giappichelli, 2021, pp. 215-232.

Per concludere la nostra analisi della tassonomia concettuale implicata nella narrazione dell'IA e nella sottesa ideologia antropomorfizzante, occorre richiamare gli ultimi due concetti *primi*, ossia responsabilità e persona. Ci si chiede se l'IA possa essere considerata responsabile delle azioni che realizza, ad esempio, mediante dispositivi robotici, e quindi se possa essere qualificata come persona, ossia agente morale.

Proprio l'accostamento alle macchine attraverso le nozioni di autonomia e intelligenza crea aspettative alte rispetto alla capacità dei sistemi intelligenti di comportarsi come persone, ossia agenti morali. Come è stato osservato «because of their increased intelligence, autonomy, and interaction capabilities, AI systems are increasingly perceived and expected to behave as moral agents»¹¹⁴. Queste aspettative dipendono in buona parte da un'intrinseca predisposizione degli esseri umani a considerare simili all'uomo macchine con certe caratteristiche, in quanto «[e]ven if you know a robot has very little autonomy, when something moves in your space and it seems it has a sense of purpose, we associate that with something having an inner awareness or goals»¹¹⁵. Il concetto di persona rappresenta un lasciapassare formidabile per l'ingresso nel mondo delle norme, regole, diritti e doveri che governano il contesto delle relazioni umane. La definizione di chi è persona e chi non lo è ha implicazioni pratiche per la distribuzione dei benefici: è una battaglia di potere. Essere persone nel senso di agenti morali implica doversi/potersi assumere anche responsabilità morale per le azioni che si compiono. Significa salire sul gradino più alto del podio, quello in cui sinora sono stati collocati solo gli esseri umani. Non è cosa di poco conto *pensare e scegliere di definire* l'IA come *persona*. Con una scelta di questo tipo, dalle forti implicazioni morali e giuridiche, si cancellano i limiti tracciabili tra noi e il frutto della nostra creatività.

Come avremo modo di vedere nel capitolo dedicato alle questioni giuridiche, in gioco vi è la stessa tenuta dei diritti fondamentali, storicamente non a caso denominati *droits de l'homme*.

In breve, alla luce delle considerazioni svolte in questo paragrafo dovrebbe essere emersa la forza della questione semantica nella determinazione dell'indi-

¹¹⁴ V. DIGNUM, *Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way*, Cham, Springer, 2019, p. 36. Non a caso è entrato nell'uso il termine *tecnologie morali* che può essere inteso come il desiderio di dare norme etiche alle macchine o per sottolineare come oggi la tecnologia sia sempre più in grado di influenzare, sebbene indirettamente, le scelte morali degli esseri umani. Sul concetto di *tecnologie morali* cfr. M. ALFANO, M. CHRISTEN, *A framework for understanding and evaluating moral technologies*, in S. SALARDI, M. SAPORITI (a cura di), *Emerging 'moral' technologies and the ethical-legal challenges of new subjectivities*, Torino, Giappichelli, 2020, pp. 1-21.

¹¹⁵ R. LERMAN, *Be wary of robot emotions; 'Simulated love is never love'*, in *AP News*, del 26 aprile 2019 accessibile al link <https://apnews.com/article/san-francisco-north-america-technology-business-ap-top-news-99c9ec8ebad242ca88178e22c7642648>.

Anche Frank Pasquale ravvisa problemi etici derivanti dal fatto che i robot suscitano sensazioni di fiducia e amicizia, cfr. F. PASQUALE, *Le nuove leggi della robotica*, cit. p. 69.

rizzo che può intraprendere o meno il rapporto tra potere e linguaggio e quanto il controllo sul linguaggio diventi un essenziale strumento di difesa o, in alternativa, di disprezzo della dignità dell'essere umano.

5. La questione semantica come problema etico

A muovere dall'analisi semantica che precede, possiamo affermare che la narrazione dell'IA, in tutte le declinazioni e sfumature richiamate, non è mai stata neutrale, ossia avulsa dal contesto socioculturale di riferimento, dalle pressioni di chi esercita il potere economico e dai valori sottesi alla cornice politico-giuridica in cui si sviluppa¹¹⁶. Anche la parte descrittiva di tale narrazione, ovvero quella effettuata da tecnici ed esperti del settore, è, a sua volta, un prodotto del contesto storico-culturale in cui il progresso scientifico ha luogo e in cui gli esperti e i tecnici si trovano ad operare. Sul punto si è ben espresso Sadin, richiamando l'attenzione sulla singolarità del racconto storico che contraddistingue l'evoluzione dell'informatica. L'autore osserva che rispetto a questa disciplina «è venuto a costituirsi una sorta di grande racconto, ripetuto in modo sempre identico, che viene a formare un quadro dall'aspetto statico»¹¹⁷, non compatibile con l'irriducibilità dell'esperienza umana, che di solito la storia racconta. Una storia, quella che accompagna l'evoluzione dell'IA, che appare del tutto decontestualizzata e che ha alimentato il mantra della neutralità della tecnica. Come però abbiamo osservato, proprio i concetti *primi* adoperati nella narrazione dell'IA sono la spia più indicativa della non neutralità di quello che avviene in questo contesto.

Se, infatti, la narrazione non può essere considerata neutrale, non è principalmente un problema di linguaggio, ma di uso del linguaggio. Trattasi di un uso strumentale effettuato da diversi attori chiave della narrazione in vista del racconto che si vuole presentare ai destinatari finali, in questo caso agli utenti, ossia ai cittadini o più in generale al pubblico. In altre parole, è una questione di scelta.

Così come vi è una narrazione che, alimentata dalla dialettica degli opposti e dall'ideologia antropomorfizzante, fa comodo a pochi detentori del potere, così un'altra narrazione è possibile ed è di fatto più plausibile»¹¹⁸. Questa è l'opinione, ad esempio, di Frank Pasquale. La sua analisi è volta a dimostrare che è possibile tracciare, attraverso scelte coraggiose, un diverso itinerario dello sviluppo robotico e dell'IA, partendo da tre presupposti: 1) sostenere la funzione complementare e non sostitutiva dell'IA rispetto al lavoro umano; 2) non pro-

¹¹⁶ Cfr. L. WINNER, *Do Artifacts have Politics?*, in *Daedalus*, 1980, 109,1, pp. 121-136.

¹¹⁷ E. SADIN, *Critica alla ragione artificiale*, cit. p. 29.

¹¹⁸ F. PASQUALE, *Le nuove leggi della robotica*, cit. p. 20.

cedere a mutamenti radicali in tutti i settori, ma mantenere lo status quo in certi casi; 3) esistono istituzioni che possono perseguire gli scopi testé elencati¹¹⁹. Queste tesi sono del tutto condivisibili e si ritiene che sia opportuno perseguirne la realizzazione non solo seguendo le proposte e i suggerimenti dell'autore, ma altresì e soprattutto attraverso il controllo sul linguaggio che proponiamo in questo lavoro. Si ritiene, infatti, che senza questo specifico riguardo per il linguaggio della narrazione, sarà più difficile sfuggire o almeno arginare utilmente le logiche imprenditoriali e commerciali che mirano a un'automazione quasi totale delle attività umane. Ciò perché il linguaggio crea assuefazione e l'assuefazione inibisce le spinte per un cambiamento di rotta.

Come vedremo di seguito, uno dei principali fattori, che ha consentito a queste strategie di insinuarsi, è la persistente e apparentemente inattaccabile visione idilliaca di un futuro di benessere e prosperità realizzabile praticamente solo tramite l'apporto tecnologico. E il linguaggio ha giocato un ruolo cruciale nella presentazione di questo progetto. Si pensi all'esordio del *Libro bianco sull'intelligenza artificiale – Un approccio europeo all'eccellenza e alla fiducia*, del 19 febbraio 2020¹²⁰, che così apre il documento «[L]'intelligenza artificiale si sta sviluppando rapidamente. Cambierà le nostre vite migliorando l'assistenza sanitaria (ad esempio rendendo le diagnosi più precise e consentendo una migliore prevenzione delle malattie), aumentando l'efficienza dell'agricoltura, contribuendo alla mitigazione dei cambiamenti climatici e all'adattamento ai medesimi, migliorando l'efficienza dei sistemi di produzione mediante la manutenzione predittiva, aumentando la sicurezza dei cittadini europei e in molti altri modi che possiamo solo iniziare a immaginare»¹²¹.

A questo elenco segue il richiamo a possibili criticità dell'impiego dell'IA – si parla di rischi potenziali, di meccanismi opachi, di discriminazioni e criminalità –, ma è già una scelta che muove in una direzione ben precisa quella di formulare con tanta enfasi gli aspetti positivi e di evidenziarli nella parte iniziale del Libro.

Ed è una scelta, quella europea, che, sebbene si stia liberando dalle iniziali influenze della dialettica degli opposti e punti effettivamente alla realizzazione di un'*AI for good*, perseguendo una strada di responsabilità e rispetto dei diritti e libertà fondamentali¹²², non si presenta ciononostante priva di persistenti incongruità e ambiguità, che sono spesso riconducibili all'uso del linguaggio.

Sottesi all'approccio istituzionale europeo ci sono, infatti, ancora frammenti

¹¹⁹ *Ibidem*, p. 19. L'autore individua la premessa generale del suo libro come segue «abbiamo mezzi per gestire le tecnologie dell'automazione e non farci imprigionare o trasformare da loro».

¹²⁰ Accessibile al sito https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_it.pdf, ultimo accesso agosto 2021.

¹²¹ *Ibidem*.

¹²² Analizzeremo nel corso del lavoro anche la proposta di Regolamento europeo relativa all'IA.

della narrazione dell'IA volti a supportare l'idea di una rivoluzione come rottura radicale con il passato e a promuovere, artatamente, immagini di una tecnologia ultimativa e risolutiva di tutti i problemi con cui attualmente si confronta l'umanità, da quelli sanitari a quelli ambientali. Tra le righe dei discorsi istituzionali sul tema si colgono ancora stralci di ipostatizzazioni oggettivizzanti e assolutizzanti dell'IA, che influenzano anche il modo di considerare i dati di cui questa si nutre. Questi ultimi vengono spesso intesi come un presupposto pre-sociale, quando invece sono essi stessi il frutto di un costruito sociale e culturale¹²³ a cui sono sottese precise scelte etiche. La visione c.d. *data-driven* della società, in cui i dati vengono usati per processi decisionali decentralizzati e più efficienti, è spesso confusa con un modello sociale in cui l'IA è in grado di dare risposte oggettive, complete e definitive sul reale e su ciò che lo compone. Ciò accade perché questa visione è contrapposta a quella c.d. *human-centric*. Ancora una volta la dialettica degli opposti si manifesta nella contrapposizione tra un'IA che sarebbe incentrata sulle persone e sulla quale l'essere umano esercita un controllo delle attività e dei suoi esiti e una che, invece, si muove come forza indipendente dall'uomo e che costruisce una propria interpretazione del reale in virtù della massa di dati di cui dispone. Nella confusione generata da questa netta contrapposizione tra uomini e macchine, diviene alquanto difficile individuare lo scarto sottile, ma cruciale, tra una società *data-driven* e una *human-centric*. Lo scarto riguarda l'ampiezza della distribuzione dei benefici del processo di digitalizzazione e trasformazione tecnologica della società. Nel caso di un approccio esplicitamente incentrato sulle persone, si sposa espressamente una dimensione egualitaria secondo cui tutti gli umani dovranno potersi avvantaggiare dello sviluppo tecnologico. Nel caso, invece, di un approccio che privilegia una visione basata sui dati e sui sistemi che macinano i dati, la distribuzione dei benefici sarà meno chiaramente ispirata al modello egualitario e contribuirà facilmente a celare i reali destinatari dei benefici delle trasformazioni tecnologiche, ossia i detentori del potere tecnologico ed economico.

Svelare la falsa dicotomia tra una società *data-driven* e una *human-centric* rappresenta il presupposto per identificare la questione etica che sta dietro la scelta semantica, ossia che la partita attorno all'IA è prima di tutto una partita di potere e di controllo del potere. Se ciò è vero, allora la questione del rigore semantico è innanzitutto «pregiudiziale alla stessa tenuta dello Stato di diritto e della democrazia»¹²⁴ ed è quindi, prima di tutto, una questione che concerne il

¹²³ D. BALAZKA, D. RODIGHIERO, *Big Data and the Little Big Bang: An Epistemological (R)evolution*, in *Frontiers in Big Data*, 2020, 3, 31, pp. 1-10, ivi, p. 3. Gli autori osservano che «[o]ne of the assumptions that allows for the objectivistic rhetoric of big data is the pre-social origin of collected data».

¹²⁴ L. FERRAJOLI, *Dei diritti e delle garanzie. Conversazione con Mauro Barberis*, Bologna, Il Mulino, 2013, p. 22.

linguaggio istituzionale, politico e legale perché ha un “carattere di preminente interesse generale”, date le possibili ricadute sui diritti fondamentali e sulla tenuta dei sistemi democratici.

A muovere dalle osservazioni che precedono, possiamo affermare che la questione semantica – o se si preferisce del rigore o controllo sul linguaggio – è, quindi, profondamente etica.

Il modello di una società in cui le risposte si trovano nei dati interpretati algoritmicamente è sostenuto dal ricorso al c.d. determinismo tecnologico. Il concetto di determinismo ricorre spesso nel dibattito sugli avanzamenti scientifici. In particolare, il contesto della genetica rappresenta un esempio paradigmatico delle discussioni attorno al c.d. determinismo genetico o essenzialismo genetico¹²⁵. Riferito più specificamente alla tecnologia «[t]echnological determinism” refers to a pervasive, yet controversial, theory about the relationship between technology and society»¹²⁶.

La nozione ha due principali accezioni, che riassumeremo brevemente di seguito – tagliando parecchio rispetto all’evoluzione del concetto per i fini d’interesse di questo lavoro, rimandando ulteriori approfondimenti alla bibliografia dedicata¹²⁷.

In una prima accezione, si ritiene che a guidare e orientare lo sviluppo sociale e i valori culturali ed etici sia lo sviluppo tecnologico che opera in maniera autonoma seguendo delle logiche interne, indipendenti dal contesto sociale. In una seconda e correlata accezione, lo sviluppo tecnologico determina il percorso della società in maniera predeterminata, in quanto la tecnologia è ritenuta sostanzialmente neutrale, ossia prescinde dal contesto in cui si realizza ed è in pratica inevitabile.

Il determinismo tecnologico nella seconda accezione confonde fatti e processi. Lo sviluppo tecnologico non è un mero fatto che si dà una volta per tutte nella realtà, è un processo mutevole e fortemente influenzato da logiche economiche e di potere. In quanto processo che si svolge nel tempo, esso risente della cultura e dei valori sociali e morali in cui è inserito. Detto diversamente, se si tiene distinto il piano descrittivo da quello prescrittivo, si può sottolineare che l’*essere*, ovvero il percorso dello sviluppo tecnologico, non porta a considerarlo

¹²⁵ Cfr. P. BORSELLINO, *Bioetica tra ‘moralì’ e diritto*, cit.; S. SALARDI, *Test genetici tra determinismo e libertà*, Torino, Giappichelli, 2010; M. BALISTRERI, G. CAPRANICO, M. GALLETI, S. ZULLO, *Bioteologie e modificazioni genetiche*, Bologna, Il Mulino, 2020.

¹²⁶ R.R. KLINE, *Technological Determinism*, in *International Encyclopedia of the Social & Behavioral Sciences* (Second edition), Amsterdam, Elsevier, 2015, pp. 109-112, ivi, p. 109.

¹²⁷ *Ibidem*, dove l’autore ripercorre la storia del concetto e la varietà dei significati stratificati con l’uso, analizzando molta della principale bibliografia dedicata al tema. Cfr. anche lo Studio del Panel for the Future of Science and Technology, intitolato *Artificial Intelligence: From ethics to policy*, 2020, p. 15 ss.

come un mero dato di fatto avulso da scelte di valore o come qualcosa di inevitabile. Non è una concatenazione causale di fatti naturali che ha prodotto lo sviluppo tecnologico, ma precise scelte umane volte a perseguire, nonostante problemi e criticità, l'obiettivo della costruzione di macchine *intelligenti*, inaugurato con il già richiamato *Dartmouth Summer Project on Artificial Intelligence* del 1956. Chi non tiene distinto il piano descrittivo da quello prescrittivo – e questo accade nella comunicazione mediatica, così come in buona parte di quella istituzionale – intenzionalmente o meno fa passare come «meri fatti constatabili empiricamente quelli che sono in realtà già espressione di valutazioni precedenti»¹²⁸, fondate su scelte culturali e valoriali. In questa prospettiva, la tendenza è quella di accettare senza domande la tecnologia, soprattutto senza domande sul potere che governa tale tecnologia¹²⁹.

La ricostruzione storica dello sviluppo tecnologico dell'IA mostra, però, come, fin dai suoi primi passi, l'IA sia stata oggetto di attenzione e progettualità finanziata da potenti attori, come la *Rockefeller Foundation*, che hanno incisivamente influenzato e indirizzato le scelte politiche di investimento. La continuità e l'irreversibilità di tale progresso – per riprendere i termini impiegati nel primo paragrafo – non sono dovuti a un *naturale* processo evolutivo, ma all'impegno umano per spingere quel processo verso specifici obiettivi¹³⁰. Ciò a dire che, se si fosse voluto rendere veramente democratica la trasformazione tecnologica si sarebbero dovuti rendere espliciti i *principi direttivi primi* su cui si sono basate *ab origine* le scelte d'indirizzo del progresso tecnologico. Ma ciò non è avvenuto. Invece, l'atteggiamento generale dei grandi della tecnologia, appoggiati spesso anche istituzionalmente, è stato per lungo tempo quello di mantenere celati questi *principi primi* attraverso una narrazione prescrittiva, camuffata da descrittiva. Il forte contributo dell'orientamento filosofico transumanista¹³¹ ha rafforzato questo camuffamento¹³². Sebbene il movimento transumanista non sia un monolite, componendosi di almeno due correnti – una mode-

¹²⁸ N. BOBBIO, *La natura delle cose*, in ID., *Giusnaturalismo e Positivismo Giuridico*, 1972, Milano, Edizioni Comunità, pp. 197-212, ivi, 204.

¹²⁹ «L'immagine della tecnologia come forza autonoma che agisce con risultati inevitabili è stata impiegata spesso nel corso dei secoli per cancellare le impronte digitali del potere e assolverlo da ogni responsabilità», S. ZUBOFF, *Il capitalismo della sorveglianza*, cit. p. 239.

¹³⁰ P. KIRCHSCHLAEGER, *Digital Transformation and Ethics*, cit. p. 60. Secondo l'autore «[e]thical discourse of technology depends on the understanding that technology is “something made” and not “anything given”».

¹³¹ P. BENANTI, *The cyborg: corpo e corporeità nell'epoca del post-umano*, Assisi, Cittadella editrice, 2016, p. 138. Secondo l'autore «[i]l movimento *transumanista* ha una particolare influenza negli ambienti della ricerca e dello sviluppo delle nuove tecnologie».

¹³² J. TESTART, A. ROUSSEAU, *Au péril de l'humain*, cit. p. 148. L'autore osserva che «[p]our les transhumanistes nourris à l'informatique, l'intelligence artificielle (IA) ou augmentée reste la première mission pour améliorer l'humanité».

rata e una *strong*¹³³ – è possibile affermare che, in generale, «il *transumanesimo* accoglie senza domande etiche particolari l'uso della tecnologia per raggiungere un *enhancement* dell'essere umano»¹³⁴, in pratica lo sostiene attraverso un implicito motto che prevede che tutto ciò che è tecnologicamente disponibile per migliorarsi vada sempre utilizzato. In questa visione, invero, i limiti e la finitezza umana sono considerati sempre e comunque dei mali da superare. Sostenendo in forma estrema questa posizione si favoriscono indirettamente argomenti a favore della prima accezione di determinismo tecnologico, richiamata in precedenza¹³⁵. Tuttavia, lungi dall'essere neutrale e in grado di autogiustificarsi, lo sviluppo tecnologico è un fenomeno che è, a sua volta, figlio di controversie e lotte morali, così come politiche e sociali «relative all'affermazione di imposizione in gruppi sociali di certe piuttosto che di altre direttive»¹³⁶. Sottolineare lo stretto legame tra potere e sviluppo tecnologico non implica affatto mettere in discussione la logica interna ai processi sperimentali che è garanzia della validità e dell'effettività delle innovazioni tecnologiche. Ma una cosa è la logica interna alla sfera tecnico-scientifica, che Scarpelli chiama *contesto di controllo*¹³⁷, altra cosa è, invece, il *contesto di invenzione*, che colloca lo sviluppo tecnologico entro una precisa cornice storica, culturale e politica. L'impiego della nozione di *contesto di invenzione*, mutuata dalla filosofia della scienza, contribuisce a comprendere i percorsi seguiti dallo sviluppo tecnologico e le scelte relative agli investimenti in un dato momento storico. Come ha chiarito Scarpelli, quello che cade in questo contesto è un insieme di «attività comportanti non oggettive preferenze ed iniziative»¹³⁸, bensì attività inserite in uno specifico contesto storico e socioculturale e come tali suscettibili come «tutte le

¹³³ Cfr. P. BENANTI, *The cyborg: corpo e corporeità nell'epoca del post-umano*, cit. p. 134 ss. Cfr. anche per un approfondimento sulle diverse sfumature entro il movimento J. TESTART, A. ROUSSEAU, *Au péril de l'humain*, cit.; F.H. LLANO ALONSO, *Homo excelsior. Los límites ético-jurídicos del transhumanesimo*, Valencia, Tirant lo Blanch, 2018. Si veda anche L. PALAZZANI, *Il potenziamento umano. Tecnoscienza, etica e diritto*, Torino, Giappichelli, 2015.

¹³⁴ P. BENANTI, *The cyborg: corpo e corporeità nell'epoca del post-umano*, cit. pp. 137-138. Cfr. anche l'articolo apparso su *Wired.it* il 29 settembre 2019 dal titolo *Mark O'Connell: "il transumanesimo crede che si possa sconfiggere la morte con la tecnologia"* in cui si sottolinea la presenza di grossi problemi etici rispetto all'uso delle tecnologie per spingere al limite la vita umana. https://www.wired.it/attualita/tech/2019/09/29/mark-o-connell-transumanesimo/?refresh_ce=.

¹³⁵ Nel dibattito i transumanisti sono «custodi [...] di una nuova libertà, quella appunto di usare senza limiti l'inedito potere di cui siamo investiti», cfr. S. RODOTÀ, *Il diritto di avere diritti*, cit. p. 355.

¹³⁶ U. SCARPELLI, *La «grande divisione» e la filosofia della politica*, cit. p. 121. L'autore fa queste osservazioni in merito al rapporto tra scienza e morale, ma sono estendibili anche al tema della tecnologia.

¹³⁷ *Ibidem*, p. 134.

¹³⁸ *Ibidem*.

attività sociali, di valutazioni morali e politiche in rapporto alla loro genesi, alle loro vicende ed ai loro esiti»¹³⁹. Se si tiene ferma questa distinzione tra *contesto di controllo* e *contesto d'invenzione*, si riesce a evidenziare che il cammino del progresso scientifico e tecnologico non è predestinato, inevitabile e irrinunciabile. Proprio la consapevolezza della non neutralità di questo progresso apre le porte a una valutazione critica delle sue direzioni e delle scelte che vengono operate, mettendo altresì in luce che la sua pervasività richiede l'irrinunciabile *participatio* non solo degli attori della produzione e commercializzazione della tecnologia, ma anche dei destinatari passivi¹⁴⁰. Se ci si muove in questa direzione, vi è un ulteriore passo che deve essere compiuto. Ci si riferisce alla necessità di rendere la partecipazione di tutti realmente effettiva¹⁴¹, non solo riconoscendo formalmente l'autonomia delle persone – come fanno molti transumanisti –, ma predisponendo le condizioni per la sua realizzazione. In altre parole, bisogna mettere nelle condizioni di essere autonome nelle scelte attraverso la messa a disposizione di strumenti di facile utilizzo per informarsi, porre domande e trovare risposte. Se si segue questa strada, ci si allontana dalle posizioni libertarie del transumanesimo, che si pronunciano a forza contro l'intervento delle istituzioni e del diritto in materia di innovazione tecnologica. Queste posizioni si fondano sul presupposto che, data l'autonomia dei destinatari delle tecnologie, bastino le regole del libero mercato a orientare le scelte. In altre parole, in questa visione l'autonomia è qualcosa di innato negli esseri umani. Mentre più realisticamente occorre riconoscere che l'autonomia è un *work in progress* durante l'esistenza di un individuo, come tale va nutrita e accompagnata nel suo sviluppo attraverso un atteggiamento proattivo da parte di chi detiene il potere informativo e comunicativo.

In sintesi, ci troviamo confrontati con una trappola linguistica – quella dell'autonomia solo formalmente affermata – alla base dello squilibrio narrativo che si esprime con eccessivo trionfalismo intorno all'IA. Come sostiene Frank Pasquale, i passi intrapresi da aziende e istituzioni pubbliche «cambiano l'equilibrio tra macchine e umani nella nostra vita quotidiana. Evitare le conseguenze

¹³⁹ *Ibidem*.

¹⁴⁰ Esistono isolate iniziative volte a instaurare un dialogo attivo tra ricercatori e società civile, che tuttavia di solito riguardano istituzioni pubbliche e non anche soggetti privati. Il caso di riferimento è l'iniziativa della Scuola Superiore Sant'Anna di Pisa, la quale ha creato una casa-laboratorio in cui si lavora a nuovi concetti di automi unitamente agli abitanti di Peccioli, cfr. l'articolo su *Il Sole 24 Ore* del 2 settembre 2021 intitolato *Il dialogo tra ricerca e persone crea robot sempre più sociali*, a firma di Riccardo OLDANI.

¹⁴¹ Secondo Peter G. Kirchsclaeger l'educazione di tutti gli attori in gioco è una questione principalmente etica e la soluzione «regarding digital transformation and data-based systems from an ethical perspective involves education embracing both the structural and systemic dimension as well as the individual dimension of the concrete actors (for example, teachers, pupils, students)», P. KIRSCHSCHLAEGER, *Digital Transformation and Ethics*, cit. p. 391.

peggiori della rivoluzione dell'intelligenza artificiale e allo stesso tempo sfruttare il suo potenziale dipenderà dalla nostra capacità di operare su questo equilibrio»¹⁴².

Gli interventi per mantenere questo equilibrio passano anche dalla narrazione che deve trovare la giusta misura tra euforia e demonizzazione. Ciò sarà possibile se si abbandonano atteggiamenti dogmatici camuffati dietro asserzioni sul progresso tecnologico apparentemente neutrali e se si rispetta la coerenza tra ciò che è realmente possibile fare nell'attuale momento storico con le tecnologie a disposizione e l'uso potenziale che, invece, viene enfatizzato, se non assolutizzato, nella narrazione mediatica tanto da confondere ciò che è con ciò che *potrà essere*¹⁴³, ma non è detto sarà.

Come ci ricorda George Orwell «il grande nemico di un linguaggio chiaro è l'insincerità»¹⁴⁴, la quale si avvale di eufemismi, argomentazioni fallaci, fumosa imprecisione per camuffare uno «scostamento tra gli scopi reali e quelli dichiarati»¹⁴⁵. Quanto affermato sinora circa l'utilità di un approccio linguistico-semantico improntato al rigore e mai dimentico della distinzione tra piano descrittivo e piano prescrittivo per smascherare posizioni ideologiche dietro l'eccessiva euforia ma anche il catastrofismo, torna altresì utile quando si sposta l'attenzione dal fronte dei transumanisti a quella dei bioconservatori.

Sebbene gli autori appartenenti a questa corrente abbiano posizioni interne con sfumature diverse, in cui qui non entreremo, vi sono alcune premesse che li accomunano. Per quello che interessa il nostro discorso, possiamo richiamare un precursore degli attuali bioconservatori, Hans Jonas. Egli scrive che «[f]ondare il "bene" o il "valore" nell'essere significa colmare il presunto divario fra essere e dover essere. Infatti, se il bene o ciò che vale è tale a partire da sé stesso e non soltanto grazie a un desiderio, un bisogno o una scelta, allora è concettualmente definibile come quella cosa la cui possibilità include l'esigenza della sua realtà, diventando così un dover essere»¹⁴⁶. Partendo dal presupposto che esiste una

¹⁴² F. PASQUALE, *Le nuove leggi della robotica*, cit. p. 19.

¹⁴³ In un articolo pubblicato su *L'Economia del Corriere della Sera* del 30 ottobre 2017 da Edoardo SEGANTINI, intitolato *Intelligenza artificiale. Basta retorica, dice Varone*, si sottolinea come «[s]entir parlare di innovazione senza retorica futurista è già inusuale». Nella maggior parte dei titoli di giornale, che trattano questa questione, vi è una parte descrittiva dell'attuale status quo, ma immediatamente accompagnato dal futuro di quello che sarà o potrà essere. Ad esempio, in *La Repubblica* del 24 aprile 2019 l'articolo di Luisa DAMIANO e Paul DEMOUCHEL intitolato *Se il badante è una macchina. I nuovi robot avranno delle funzioni socialmente significative*; in *La Repubblica Robinson* del 7 gennaio 2018 l'articolo a firma di Jaime D'ALESSANDRO intitolato *Cara AI impara dai bambini. Le previsioni di Yann LeCun, il guru di Facebook: faremo evolvere le macchine insegnando loro a pensare come noi*.

¹⁴⁴ G. ORWELL, *La politica e la lingua inglese*, in S. PERRELLA (a cura di), *Nel ventre della balena*, Milano, Bompiani, 2012.

¹⁴⁵ *Ibidem*.

¹⁴⁶ H. JONAS, *Il principio di responsabilità. Un'etica per la civiltà tecnologica*, Torino, Einaudi, 1990, p. 101.

natura umana che segue l'ordine naturale delle cose, si ritiene che essa vada preservata e conservata sempre e comunque. In questa prospettiva, la tecnologia è considerata quasi esclusivamente per il suo potenziale nocivo, in quanto è mezzo per allontanarsi dallo stato naturale originario, laddove gli scopi perseguiti non siano quelli di mera conservazione/cura dello status quo¹⁴⁷.

Ecco perché i bioconservatori sono a favore di un uso della tecnologia esclusivamente a fini curativo/terapeutici e sono invece a sfavore di impieghi *oltre la terapia*. Anche queste posizioni, condividono con quelle transumaniste, seppur nella diversità che le caratterizza, una concezione ideologica, affetta da dogmatismo, oggettivismo e ontologismo. Le ricadute di questo approccio sono dirette a limitare l'uso delle tecnologie in maniera tanto più spinta, quanto più è il loro potenziale invasivo. Ciò perché si dà per scontata l'esistenza di un bene ultimo che si realizza in un «ideale di una vita buona e virtuosa»¹⁴⁸.

Dalle considerazioni che precedono dovrebbe essere emersa l'insostenibilità di posizioni che sostengono tesi estreme. Non si tratta, infatti, di demonizzare o divinizzare il macchinico, ma di poter compiere scelte rispetto a questo fenomeno che siano improntate ai valori «della libertà e della trasparenza delle scelte»¹⁴⁹, in quanto sostenute da «una volontà etica di consapevolezza etica»¹⁵⁰ per realizzare il benessere generale.

6. La questione semantica come questione costituzionale e democratica

Per una *participatio* alla trasformazione digitale della società che si possa considerare effettiva, attiva e produttiva per tutti, occorre che l'informazione che circola abbia alcuni requisiti, che in parte sono già emersi nei paragrafi precedenti. Non solo tale informazione deve essere veritiera, in particolare deve esserci corrispondenza e coerenza tra ciò che è e ciò che viene descritto, deve altresì essere chiara, comprensibile, *formativa* ma non *manipolativa*. Per avere queste caratteristiche occorre che nei contesti in cui si narra pubblicamente di IA si ponga attenzione alla questione semantica, ossia al rigoroso controllo linguistico e alla distinzione tra piano prescrittivo e piano descrittivo. Non è una richiesta che può essere investita di una mera valenza soggettiva, bensì ha una portata più ampia e chiama in causa la dimensione democratica e costitu-

¹⁴⁷ Cfr. T.L. EISSA, *Gesünder, intelligenter, perfekt? Selbstgestaltung durch Enhancement im Kontext pluralistischer Ethik*, Freiburg, Verlag Karl Alber, 2014.

¹⁴⁸ P. BORSELLINO, *Bioetica tra 'moralì' e diritto*, cit. p. 34.

¹⁴⁹ P. BORSELLINO, *Morale, diritto e positivismo giuridico nel pensiero di Uberto Scarpelli*, in *Bioetica tra autonomia e diritto*, Milano, Zadig, 1999, pp. 213-276, ivi, p. 238.

¹⁵⁰ U. SCARPELLI, *Etica, linguaggio, ragione*, in ID., *L'etica senza verità*, Bologna, Il Mulino, 1982 pp. 49-72, ivi, p. 68.

zionale del processo comunicativo-informativo. Come un tempo si riteneva che l'informazione mediante la televisione pubblica, dovesse rispettare alcuni canoni di correttezza, buona fede e pluralismo¹⁵¹, così oggi occorre che i nuovi luoghi in cui si forma il senso comune e il consenso e dove si esercita un penetrante controllo sul dibattito pubblico, sugli umori e sugli orientamenti dei consociati, ad esempio i social media, rispettino almeno gli stessi obblighi di corretta informazione. Come è noto, vi è una stringente sinergia tra potere economico, tecnologico e mediatico che esercita una importante influenza sulla formazione dell'opinione pubblica¹⁵². Il linguaggio diviene, quindi, uno strumento basilare per (dis)orientare le condotte dei destinatari dell'informazione. Il controllo sul linguaggio attraverso il rispetto del rigore semantico diviene, quindi, una questione costituzionale e democratica, che passa attraverso la ridefinizione del rapporto tra media tradizionali e nuovi media. La diversità tra questi mezzi di comunicazione non deve rappresentare una scusa per non garantire una buona informazione. In altre parole, le caratteristiche peculiari dei nuovi media non implicano necessariamente la conclusione che i doveri posti in capo ai giornalisti che usano mezzi tradizionali non debbano trovare equivalenza nel nuovo mondo informativo.

Il dovere di accuratezza delle notizie imposto ai media tradizionali, che passa anche dall'uso del controllo sul linguaggio, può e deve essere imposto anche a quelli nuovi¹⁵³. Ciò perché il mancato controllo di veridicità e attendibilità delle notizie consente di occultare i (dis)valori che sono sottesi alla scelta di (dis)informazione, a mezzo della quale si mantengono le asimmetrie funzionali a esercitare occultamente forme di potere e controllo sugli individui. In questo scenario, «si torna così alla necessità di non perdere mai di vista quella costituzionalizzazione della persona che è il risultato di un lungo processo»¹⁵⁴, che ha fatto del libero sviluppo della propria personalità un diritto fondamentale. Un

¹⁵¹ Come nel campo dell'informazione a mezzo televisivo risulta deleterio per la democrazia che ci sia un solo padrone che detiene anche potere economico e politico, altrettanto pericoloso è «che una sola azienda raccolga tanti dati e abbia tanto potere sugli utenti», cfr. F. PASQUALE, *Le nuove leggi della robotica*, cit. p. 121, dove l'autore mette in evidenza che il potere congiunto di Instagram, Facebook e WhatsApp nelle mani di Mark Zuckerberg consente di mettere in atto manipolazioni e ingerenze sul piano pubblico come quelle perpetrate durante le elezioni americane.

¹⁵² Cfr. sul fenomeno della disinformazione L. MORLINO, M. SORICE (a cura di), *L'illusione della scelta. Come si manipola l'opinione pubblica in Italia*, Roma, LUISS University Press, 2021.

¹⁵³ Ricordiamo in questa sede, ma non ci addentreremo nello specifico della questione, il ruolo che possono giocare, se bene utilizzate le tecnologie per trovare e rimuovere contenuti illegali o indesiderati. Sul tema cfr. PANEL FOR THE FUTURE OF SCIENCE AND TECHNOLOGY, *Regulating disinformation with artificial intelligence*, 2019, accessibile al link <https://op.europa.eu/it/publication-detail/-/publication/b8722bec-81be-11e9-9f05-01aa75ed71a1>, ultimo accesso Agosto 2021.

¹⁵⁴ S. RODOTÀ, *Il diritto di avere diritti*, cit. p. 360.

diritto che può essere davvero realizzato solo se la libertà di parola e di espressione non sono strumentalmente fraintesi.

Come sottolinea Frank Pasquale «da tempo i media hanno il problema della propaganda di parte, ma una sfera pubblica prevalentemente automatizzata può aggravarlo, consentendo a bugie e montature di diffondersi in modo virale»¹⁵⁵. Come garantire che i sistemi automatizzati rispettino quei doveri di «agire in buona fede e fornire informazioni precise su questioni di interesse generale, sulla base di fatti accertati con ricorso a fonti credibili»¹⁵⁶ che sono imposti ai giornalisti umani? Eppure, il rispetto di questi doveri è fondamentale per consentire ai destinatari dell'informazione di comprendere ciò che sta accadendo attorno a loro e rendersi conto di eventuali violazioni di diritti di cui sono titolari e avere a disposizione gli strumenti per agire in giudizio per la riparazione di tali violazioni. Basta pensare al complesso tema della protezione dei dati. Il Regolamento europeo 679/2016 sulla protezione dei dati (GDPR) garantisce molti diritti al soggetto interessato dal trattamento dei dati, ma non è sempre facile per chi non è del mestiere rendersi conto che può esserci in atto una violazione di qualcuno di questi diritti. Questo è un problema culturale, sociale, informativo e linguistico.

Il diritto alla libertà di informazione implica un diritto negativo alla non disinformazione¹⁵⁷, che discende dalla libertà di coscienza e di pensiero, ossia la prima «libertà fondamentale che si è affermata nella storia del liberalismo e che implica il diritto alla non manomissione della propria coscienza provocata da disinformazione intorno ai fatti e alle questioni di pubblico interesse»¹⁵⁸. Anche perché la disinformazione su materie di interesse pubblico mina alla base altri diritti fondamentali come quello all'autonomia e al libero sviluppo della propria personalità. Pertanto, se riconosciamo, come ha fatto la Corte europea per i Diritti Umani, l'importanza crescente dei social media per esercitare il diritto alla libertà di espressione e, altresì, il diritto di accedere alle informazioni¹⁵⁹, non possiamo non ritenere che vi sia, anche in questo nuovo contesto, un correlato dirit-

¹⁵⁵ F. PASQUALE, *Le nuove leggi della robotica*, cit. p. 37. Il problema della viralità è considerato «estremamente pericoloso, profondamente impattante sulla corretta formazione dell'opinione pubblica», cfr. A. SIMONCINI, *Forum Law and Artificial Intelligence*, in *BioLaw Journal-Rivista di Biodiritto*, 2020, 1, pp. 498-501.

¹⁵⁶ V. ZAGREBELSKY, R. CHENAL, L. TOMASI, *Manuale dei diritti fondamentali in Europa*, Bologna, Il Mulino, 2016, p. 316.

¹⁵⁷ L. FERRAJOLI, *Poteri selvaggi. La crisi della democrazia italiana*, Roma-Napoli, Laterza, 2011, p. 56. L'autore richiama l'articolo 19 della Dichiarazione universale dei diritti del 1948 e l'articolo 19 del Patto sui diritti civili e politici del 1966.

¹⁵⁸ *Ibidem*.

¹⁵⁹ Cfr. CEDU, Ahmet Yildirim v. Turkey, disponibile qui, <http://hudoc.echr.coe.int/eng-press?i=003-4191041-4965778>; Delfi AS v. Estonia, para 110, disponibile qui: <http://hudoc.echr.coe.int/eng?i=001-155105>.

to alla non disinformazione come poc'anzi richiamato. Ciò a maggior ragione se condividiamo la definizione di “digital citizen” come «a person who masters the competences of democratic culture in order to be able to competently and positively engage with evolving digital technologies; participate actively, continuously and responsibly in social and civic activities; be involved in a process of life-long learning (in formal, informal and non-formal settings) and be committed to defending continuously human rights and dignity»¹⁶⁰.

Inoltre, se l'informazione riguarda fatti o questioni di interesse pubblico, affinché non si trasformi in disinformazione, il rispetto, al più alto grado possibile, del controllo semantico è un prerequisito imprescindibile per una effettiva partecipazione democratica¹⁶¹.

7. Trustworthy AI: il controllo linguistico come presupposto per costruire un rapporto di fiducia con le tecnologie avanzate

Nel contesto europeo, da più di un decennio, da quando cioè si è iniziata a costruire l'intelaiatura strategica utile a governare lo sviluppo delle tecnologie intelligenti, uno dei principali concetti attorno a cui si è lavorato è quello di *fiducia*. Il riferimento passa attraverso formulazioni diverse. Si parla, ad esempio, di un'IA come *force for good in society*¹⁶², una *trustworthy AI*¹⁶³, una *AI4people*¹⁶⁴, una *good AI society*¹⁶⁵. Si insiste sul fatto che, proprio le ampie ricadute dell'IA, richiedono “fiducia dalla società” e una modalità di introduzione a livello sociale che consenta di costruire fiducia e comprensione¹⁶⁶. Nel-

¹⁶⁰ Punto 2 dell'Appendix to Recommendation CM/Rec(2019)10 of the Committee of Ministers to member States on developing and promoting digital citizenship education, accessibile al link https://search.coe.int/cm/Pages/result_details.aspx?OBJECTID=090000168098de08, ultimo accesso Agosto 2021.

¹⁶¹ S. SALARDI, *Democrazia e nuove tecnologie: scenari passati e dell'avvenire*, in Ordines, A. LO GIUDICE, T. CASADEI (a cura di), *Forum Democrazia e Parlamentarismo*, 2020, 2, pp. 397-412.

¹⁶² Cfr. *A European Approach to Artificial Intelligence*, <https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence>, ultimo accesso settembre 2021.

¹⁶³ *Ibidem*. Per un'analisi delle *Ethics Guidelines for Trustworthy AI*, cfr. M. CATANZARITI, *Etica “artificiale”: un nuovo modello regolatorio?*, in *Ars Interpretandi*, 2021, 1, pp. 165-179.

¹⁶⁴ Cfr. *AI4PEOPLE'S ethical framework for a good AI society: opportunities, risks, principles and recommendations*, <https://www.eismd.eu/ai4people/>, ultimo accesso settembre 2021.

¹⁶⁵ *Ibidem*.

¹⁶⁶ Così, ad esempio, PANEL FOR THE FUTURE OF SCIENCE AND TECHNOLOGY, *The ethics of artificial intelligence: Issues and initiatives*, 2020, p. II, consultabile al seguente link [https://www.europarl.europa.eu/tbinktank/en/document.html?reference=EPRS_STU\(2020\)634452](https://www.europarl.europa.eu/tbinktank/en/document.html?reference=EPRS_STU(2020)634452), ultimo accesso settembre 2021. Il concetto di *trust* ricorre in diversi studi del Panel citato, cfr. anche *Understanding algorithmic decision-making: Opportunities and challenges*, 2019 e *A governance framework for*

la relazione introduttiva alla Proposta del nuovo Regolamento sull'IA da parte della Commissione europea (Legge sull'Intelligenza Artificiale), si richiama, con insistenza, l'attenzione verso "un approccio europeo all'eccellenza e alla fiducia", volto a "sviluppare un ecosistema di fiducia proponendo un quadro giuridico per un'IA affidabile". In aggiunta, si sottolinea che la Proposta si "prefigge di dare alle persone e agli altri utenti la fiducia per adottare le soluzioni basate sull'IA"¹⁶⁷. Il richiamo costante e insistito su questo concetto ha la dichiarata intenzione di caratterizzare l'approccio europeo, diversificandolo da altri approcci. L'idea di una *trustworthy AI* è, almeno in linea generale, coerente con il ripetuto richiamo al rispetto dei diritti fondamentali e delle libertà dei cittadini europei nel percorso della trasformazione digitale della società europea. Tuttavia, qualche rilievo critico all'uso di questa nozione non può essere risparmiato.

I rilievi critici alla nozione di fiducia si ricollegano giocoforza alle critiche già espresse in questo lavoro all'ideologia antropomorfizzante e alla dialettica degli opposti. Sebbene la nozione di fiducia abbia iniziato ad essere studiata come proprietà delle relazioni interpersonali basate sul supporto e la cooperazione reciproca solo a partire dagli anni Cinquanta del secolo scorso¹⁶⁸, è oggi innegabile che, quando si usa questa nozione nel linguaggio comune, essa si riferisce a uno speciale legame che si instaura tra gli attori della relazione e sta a indicare che questi attori sono necessariamente persone¹⁶⁹. La nozione di fiducia (*trust*) rappresenta il perno attorno al quale si svolgono le relazioni umane. Possiamo affermare che il suo significato chiama in causa aspetti relazionali e richiede, quindi, l'esistenza di una relazione. Non è la relazione in sé, ma una sua proprietà¹⁷⁰. Si pensi alla relazione di cura tra medico e paziente, notoriamente fondata sulla fiducia reciproca. In questo specifico campo, il paziente ha fiducia nel medico non solo per le sue capacità e competenze tecniche, ma anche perché ha l'aspettativa che i suoi interessi e bisogni verranno rispettati e soddisfatti. La relazione fondata sulla fiducia si basa su credenze e

algorithmic accountability and transparency, 2019, entrambi consultabili al link <https://www.europarl.europa.eu/thinktank/en/search.html?word=Understanding+algorithmic+decision-making>, ultimo accesso settembre 2021. Cfr. anche HIGH-LEVEL EXPERT GROUP ON ARTIFICIAL INTELLIGENCE, *Ethics Guidelines for Trustworthy AI*, 2019, consultabile al link <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>, ultimo accesso settembre 2021.

¹⁶⁷ Proposta di Regolamento del Parlamento e del Consiglio che stabilisce regole armonizzate sull'Intelligenza Artificiale (Legge sull'Artificiale) e modifica alcuni atti legislativi dell'Unione, COM(2021) 206 finale, p. 1.

¹⁶⁸ J. BARBALET, *The experience of Trust: Its Content and Basis*, in M. SAZAKI (eds.), *Trust in Contemporary Society*, USA, BRILL, 2019, pp. 11-30.

¹⁶⁹ *Ibidem*, p. 14. L'autore scrive «[t]hose who give trust and those who receive trust are necessarily persons».

¹⁷⁰ M. TADDEO, *Trusting Digital Technologies Correctly*, in *Minds and Machines*, 2017, 27, pp. 565-568.

sentimenti, ma non certezze, che una persona si comporterà in modo da onorare la fiducia riposta. È una proprietà della relazione che la proietta nel futuro, fondandola però su una continuità di comportamento che richiede di essere confermata nel tempo.

Riferire la nozione di fiducia alle tecnologie può pertanto risultare fuorviante, in quanto mancano molti degli elementi richiamati oppure questi elementi funzionano diversamente. Manca l'elemento di impegno, di disposizione a tenere un dato comportamento, in quanto l'IA è (pre)programmata a tenere quel comportamento e non sviluppa una disposizione a mantenerlo nel tempo, nel senso che non può fare diversamente, a meno che non sia programmata in tal senso. Manca anche la proiezione nel futuro che la fiducia relazionale richiede, perché quando la fiducia è rivolta alle tecnologie, a contare è l'azione contingente che consente di affrontare un dato problema nel presente.

Se è vero che nel linguaggio comune la nozione può essere alternativamente riferita agli esseri umani o a prodotti tecnologici – quella che Giddens denomina *trust in abstract systems*¹⁷¹ –, va tuttavia sottolineato che la fiducia di cui si parla in relazione all'IA ha valenza metaforica¹⁷². Riferita ai sistemi intelligenti, la fiducia si traduce nell'affidabilità tecnica della loro prestazione. Rispetto alla fiducia tra esseri umani, quello che difetta, quando la nozione è riferita al macchinico, è la reciprocità relazionale, nonché lo sforzo richiesto per guadagnarsi tale fiducia attraverso la continuità dei comportamenti¹⁷³.

Quale è l'aspetto problematico dell'uso del termine fiducia riferito all'IA?

Parlare di fiducia nei confronti di sistemi artificiali o tecnologie significa «creating “access points”»¹⁷⁴, ossia punti di accesso alla fiducia, o meglio a quel significato del termine strettamente connesso all'instaurarsi di legami forti basati su credenze e sentimenti, che incentivano un atteggiamento, spesso inconscio, di umanizzazione del prodotto a cui attribuiamo fiducia¹⁷⁵. Porre l'accento sull'uso mistificante di questa nozione nella Proposta di Regolamento, non significa ritenere che il generale orientamento istituzionale europeo, formalmente volto alla tutela dei diritti fondamentali, non sia condivisibile. Quello

¹⁷¹ A. GIDDENS, *The consequences of modernity*, Oxford, Polity, 1990, pp. 83-88.

¹⁷² T. SCANLON, *Moral Dimensions. Permissibility, Meaning, Blame*, Cambridge, Harvard University Press, 2008, p. 161.

¹⁷³ P.G. KIRCHSCHLAEGER, *Digital Transformation and Ethics*, cit. p. 413.

¹⁷⁴ *Ibidem*, p. 412.

¹⁷⁵ Thomas Metzinger, membro del gruppo di esperti che ha elaborato le linee guida *Ethics Guidelines for Trustworthy AI*, 2019 cit. ha aspramente criticato il concetto utilizzato. Secondo l'autore «Artificial Intelligence cannot be trustworthy. The Trustworthy AI story is a marketing narrative invented by industry, a bedtime story for tomorrow's customers», cfr. T. METZINGER, *Ethics washing made in Europe*, in *Tagesspiegel*, 8 aprile 2019, accessibile al link <https://www.tagesspiegel.de/politik/eu-guidelines-ethics-washing-made-in-europe/24195496.html>, ultimo accesso settembre 2021.

che interessa, con la sottolineatura critica appena operata, è impedire che per altre vie, in particolare attraverso l'uso del linguaggio, questo nobile obiettivo venga depotenziato.

Un ulteriore aspetto che va considerato, quando si parla di fiducia, è la possibile confusione tra la fiducia e un atto di fede. Si dice, ad esempio, "avere cieca fiducia in qualcuno". Questa commistione di significati circola nel linguaggio ordinario. La fiducia cieca sta a indicare un'adesione a ciò che viene proposto senza che la proposta venga messa in discussione, perché ci sono condizioni tali che rassicurano chi ripone questa fiducia (es. la presenza di un soggetto esperto) e lo fanno astenere da ulteriori controlli. In questa sfumatura semantica di fiducia è implicita la presenza di un'autorità a cui affidarsi, senza mettere in discussione i suoi precetti e le prescrizioni. Sebbene sottile, la piega semantica qui evidenziata ha ricadute importanti sulle persone, se i confini, entro cui la narrazione dell'IA avviene, sono quelli antropomorfizzanti alimentati dalla dialettica degli opposti. Se si segue il racconto di un'IA neutrale, oggettiva, onnisciente è giocoforza essere predisposti a una deferenza nei suoi confronti come se fosse un'autorità. Non è richiesto che ci si ponga domande né che si sviluppi un pensiero critico nei suoi confronti. L'oracolo algoritmico ci dice tutto perché sa tutto e non deve spiegazioni sui suoi precetti. Se si è propensi a veder l'IA come un'autorità, il passaggio dall'IA che fa previsioni su attività ed eventi a un'IA che giudica gli esseri umani¹⁷⁶ sarà quasi impercettibile. Sarà difficile comprendere che «gli algoritmi sono opinioni inserite in codici. Non sono obiettivi»¹⁷⁷.

Se non si tiene alta l'attenzione su questo punto, sarà proprio la nozione di fiducia a inficiare gli sforzi europei verso un progresso incentrato sulle persone. Tale nozione si inserisce, invero, nella tassonomia concettuale della narrazione dell'IA e, con la valenza persuasiva che le è propria, rischia di inibire un approccio critico-razionale allo sviluppo tecnologico. Infatti, se si ha fiducia in qualcuno, vuole dire che ci si fida e spesso questa predisposizione d'animo segue l'atteggiamento di chi non ritiene di dover più controllare. Ma questa è un'altra trappola linguistica che dovremmo evitare.

Se si lascia troppo margine operativo alla valenza persuasiva delle nozioni come fiducia si rischia di svuotare il ruolo dello stesso diritto nell'orientare lo sviluppo tecnologico. Chi si fida non controlla e chi non è controllato non ha nemmeno bisogno di regole entro cui operare.

Come avremo modo di vedere nei capitoli che seguono, se è vero che è molto difficile mantenere sempre trasparente, semanticamente pulito ed equilibrato il

¹⁷⁶ F. PASQUALE, *Le nuove leggi della robotica*, cit. capitolo 5.

¹⁷⁷ La traduzione in italiano è dell'autrice, l'originale si trova in Y. DEMUTH, *Die unheimliche Macht der Algorithmen*, in *Beobachter*, 26 aprile 2018, la frase originale è «Algorithmen sind Meinungen in Codes eingebettet. Sie sind nicht objektiv» accessibile al link <https://www.beobachter.ch/digital/multimedia/big-data-die-unheimliche-macht-der-algorithmen>, ultimo accesso settembre 2021.

discorso sull'IA, non si deve tuttavia dedurre da questa difficoltà, l'impossibilità di un impegno per raggiungere questo obiettivo al più alto grado possibile, soprattutto quando in gioco vi sono categorie e istituti giuridici. Se vogliamo essere noi i protagonisti della profonda innovazione che interessa la società in cui viviamo, allora «[s]iamo noi cittadini che dobbiamo decidere come distribuire i benefici e controllare i rischi legati alla diffusione dell'Intelligenza Artificiale. Dobbiamo fare in modo che tutti abbiano consapevolezza e condividano i pericoli di concentrazione monopolistica, di imprecisione e del perseguimento di fini malvagi»¹⁷⁸.

¹⁷⁸ M. CHIRIATTI, *Incoscienza artificiale. Come fanno le macchine a prevedere per noi*, Roma, LUISS University Press, 2021, p. 29.

CAPITOLO II

PROGRESSO TECNOLOGICO E QUESTIONE TECNICA: LA SEMANTICA DEL *POSSE*

SOMMARIO: 1. Introduzione. – 2. Breve cronistoria dell'IA. – 3. Perché i problemi tecnici sono problemi etici? – 4. La conoscenza nella prospettiva tecnologica e in quella scientifica. – 5. La questione tecnica e la semantica dell'antropomorfizzazione: dominio e potere sull'individuo. – 6. Il problema tecnico che dissimula un problema socioculturale.

1. *Introduzione*

Nel capitolo che precede ci siamo soffermati sull'analisi dei concetti e degli argomenti che hanno caratterizzato sino ai nostri giorni la narrazione dell'IA al grande pubblico, ma che si sono altresì insinuati a livelli istituzionali europei. Come osservato, la narrazione si è sviluppata attorno a una dialettica degli opposti che si alimenta dell'ideologia antropomorfizzante. Gli effetti di tale modalità narrativa non si spiegano solo sul piano normativo con riguardo ai diritti fondamentali e ai valori a essi sottesi, ma si fanno sentire anche sugli aspetti tecnici dello sviluppo tecnologico. Il progresso tecnologico rischia di essere frainteso ed equivocato nella sua portata a causa di una narrazione ambigua. Ecco perché appare opportuno, in questo capitolo, approfondire la *semantica* delle questioni tecniche. A muovere da un breve richiamo alle tappe centrali dello sviluppo delle tecnologie intelligenti nel secolo scorso, porteremo l'attenzione sulla distinzione tra problema tecnico e problema etico. In questo modo sarà possibile far emergere le criticità correlate all'idea che parlare di IA significhi principalmente, se non esclusivamente, affrontare decisioni tecniche di esclusiva spettanza degli esperti del settore. Dopo avere chiarito quali trappole si nascondono dietro questa concezione, si analizzerà il tema della conoscenza del reale, che ha una lunga storia nel pensiero filosofico occidentale. Si cercherà di mettere in luce alcuni problemi di fondo derivanti dalla diffusa e radicata idea che tramite l'IA, in grado di macinare dati in quantità inestimabili, saremmo oggi in grado di dare una rappresentazione esauriente e integrale del reale. Dall'analisi emergerà che chi sostiene la *retorica oggettivistica* dei *big data* cela una precisa posi-

zione etica, secondo la quale l'innovazione tecnologica è neutrale rispetto a scelte di valore e a orientamenti socioculturali.

2. Breve cronistoria dell'IA

Vi sono autori che, nel raccontare l'evoluzione storica dell'IA, risalgono sino al dodicesimo secolo. Si richiama alla memoria la letteratura di quel secolo, nella quale si narra di una creatura denominata *golem*, con caratteristiche estetiche simil-uomo, ma molto più forte¹. Sebbene le origini del racconto possano essere così risalenti, per i fini del nostro argomentare, è opportuno circoscrivere la nascita dell'IA al secolo XX, entro i confini degli avanzamenti resi possibili dall'inedita disponibilità di conoscenze scientifiche e dalla possibilità di conversione tecnica di tali conoscenze. È, infatti, in questo specifico arco di tempo che l'essere umano sviluppa, grazie alla tecnologia, un potere inedito di controllo sull'esistente. Ed è in relazione a questa specifica eccedenza che si pongono problemi etici circa l'insufficienza degli strumenti elaborati nelle riflessioni svolte nell'ambito delle etiche tradizionali, principalmente, se non esclusivamente, focalizzate sull'agire nel presente. Quello che le tecnologie intelligenti mettono a disposizione dell'essere umano è la mutata capacità di agire predittivamente, quindi, con ricadute sul futuro, ed è questo che rende la riflessione attuale sulla tecnologia radicalmente diversa da quella operata in passato².

Il termine Intelligenza Artificiale (IA) fu coniato nel 1956 da John McCarthy. Mc Carthy la definì come la scienza e l'ingegneria che produce macchine intelligenti, in particolare programmi informatici intelligenti. Il percorso che ha portato agli odierni sviluppi e applicazioni dell'IA non è stato continuo. Si è di fatto costellato di c.d. inverni e primavere. Tali espressioni sono utilizzate per indicare i periodi di arresto non solo del progresso, ma delle stesse ricerche sull'IA, seguiti da periodi di forte accelerazione³. Dopo il susseguirsi di una serie di sforzi tra il 1975 e il 1985, alcuni di successo ed altri fallimentari, sarà a partire

¹ S. HADDADIN, D. KNOBBE, *Robotics and Artificial Intelligence*, in M. EBERS, S. NAVAS (eds.), *Algorithms and Law*, Cambridge, Cambridge University Press, 2020, pp. 1-36, ivi, p. 1 ss.

² H. JONAS, *Il principio di responsabilità. Un'etica per la civiltà tecnologica*, Torino, Einaudi, 2002.

³ Per un approfondimento cfr. T. CASADEI, S. PIETROPAOLI, *Intelligenza artificiale: fonte o confine del diritto?*, in ID. (a cura di), *Diritto e Tecnologie Informatiche*, Milano, Wolters Kluwer, 2021, pp. 219-232. Chiriatti critica questo approccio, sottolineando la necessità di *destagionalizzare* l'IA, perché vari sono gli aspetti che hanno contribuito a inverni e primavere, non solo la creatività umana, ma presenza o meno di finanziamenti e interesse, cfr. M. CHIRIATTI, *Incoscienza artificiale. Come fanno le macchine a prevedere per noi*, Roma, LUISS University Press, 2021, p. 67.

dal 1995 che si tornerà a investire nuovamente sull'IA⁴. Grazie all'avvento di Internet, che richiede(va) applicazioni informatiche intelligenti e grandi volumi di informazioni, si aprono le porte a un ampio impiego dell'IA in svariati ambiti e attraverso diversi strumenti e dispositivi, come i motori di ricerca, la costruzione di siti web, le transazioni commerciali ecc. Dagli anni '90 del secolo scorso in poi, molti saranno i momenti di successo registrati dall'IA. Si ricorda che «[n]el 1997, il computer Deep Blue batte il campione del mondo di scacchi, Garri Kasparov. Nel 2011, il sistema esperto di IBM, Watson, batte gli umani nel quiz televisivo Jeopardy. Dal 2015, l'IA realizza in pochi minuti analisi oncologiche che richiederebbero decenni a degli oncologi in carne e ossa»⁵. Questi successi sono basati su programmi tradizionali, con algoritmi programmati manualmente. Con queste caratteristiche l'IA riesce a svolgere in modo eccellente singoli compiti, che richiedono un'*expertise* specifica. Tuttavia, la gestione di problematiche complesse, che richiedono più abilità per il *problem-solving*, rimane, ancora per un po', fuori dalla portata di queste tecnologie. Le cose iniziano a cambiare con gli sviluppi, a partire dal 2012, del *Deep Learning*. Trattasi di una sottocategoria di *Machine Learning*⁶ (ML). A differenza di quest'ultimo, che si caratterizza per l'apprendimento supervisionato, il *Deep Learning* si basa invece su reti neurali che fanno uso di dati non strutturati e in quantità molto superiore. Le reti funzionano in modo simile al cervello umano e cercano di individuare, autonomamente dall'intervento umano, le caratteristiche distintive per la soluzione di un problema. È l'era delle macchine che autoapprendono. Un primo esempio di *Deep Learning* non supervisionato è del 2012: *Google Brain*. Quest'ultimo ha identificato autonomamente il concetto di gatto, dopo avere analizzato per alcuni giorni milioni di immagini non etichettate. A fronte di questo importante momento di svolta nell'evoluzione dell'IA si sostiene che ormai «i computer vengono educati più che programmati»⁷. Con l'uso del termine “edu-

⁴ Agli inizi degli anni '90 del secolo scorso viene coniato il termine *Singularity*, che diverrà popolare con il lavoro di Kurzweil nel 2005. Il termine indica un punto nel tempo oltre il quale lo sviluppo dell'IA diverrà irreversibile e incontrollabile, con conseguenze sconosciute per l'umanità, cfr. J. LANDGREBE, B. SMITH, *Why Machines Will Never Rule the World*, New York, Routledge, 2023, p. XII.

⁵ L. ALEXANDRE, *La guerra delle intelligenze. Intelligenza artificiale contro intelligenza umana*, Torino, EDT, 2018, p. 7.

⁶ «Over years, ML has developed into a number of different directions. By and large, they can be classified into three broad categories, depending on their learning pattern: supervised, unsupervised, and reinforcement learning», M. EBERS, *Regulating AI and Robotics. Ethical and Legal Challenges*, in M. EBERS, S. NAVAS (eds.), *Algorithms and Law*, Cambridge, Cambridge University Press, 2020, pp. 37-99, ivi, p. 43. Sui vari tipi di modelli cfr. anche M. IENCA, *Intelligenza2*, Torino, Rosenberg & Sellier, 2019, p. 85.

⁷ L. ALEXANDRE, *La guerra delle intelligenze. Intelligenza artificiale contro intelligenza umana*, cit. p. 12.

care” si richiama un ambito di attività da sempre riferite all’essere umano, tanto che per indicare attività simil-educative riferite ad altri enti come gli animali è di norma usato il termine di addestramento. Nel caso dell’IA è invece diventato usuale parlare di educare le macchine, anche se il termine corretto sarebbe addestrare (*training*).

Nel linguaggio comune, invero, il verbo “educare” richiama diverse sfumature di significato e si riferisce a un processo complesso in cui interagiscono molti aspetti. Si educa chi può comprendere, chi può maturare sul piano intellettuale e morale, chi può affinare certe capacità, non solo chi potrà poi eseguire compiti programmati da altri. Questo è un ulteriore esempio, che si aggiunge ai tanti già fatti, di come l’ideologia antropomorfizzante sia radicata nel linguaggio comune.

Non è tuttavia esclusivamente attraverso il linguaggio che si inducono fraintendimenti. Un’altra importante fonte di criticità, che può confondere anche gli esperti, è la non corretta distinzione tra tipologie di problemi che l’IA può sollevare. Mi riferisco specificamente alla distinzione tra problema tecnico e problema etico, che è ben nota nella riflessione bioetica e in quella di etica ambientale. A questa rilevante distinzione è dedicato il paragrafo che segue.

3. Perché i problemi tecnici sono problemi etici?

Da quando la trasformazione digitale della società è diventata una realtà praticabile ed è possibile estendere l’utilizzo delle tecnologie intelligenti a quasi tutte le attività umane, sono emerse non solo le potenzialità e i vantaggi di tali operazioni, ma anche una serie di criticità che chiamano in causa questioni politiche, sociali ed etiche. È su queste ultime che ci soffermeremo, in quanto toccano da vicino l’individuo, i suoi diritti e i suoi doveri.

Nel dibattito istituzionale, così come in ambito mediatico, si sente e si legge spesso la locuzione “problemi/questioni etici/etiche”. Questa espressione appare spesso usata dando per scontato che chiunque la senta sappia di cosa si sta parlando. In realtà, al di fuori del contesto della riflessione etica, non è sempre chiaro, a chi sente questa espressione, a che cosa ci si riferisca esattamente con “problemi etici”. Va, innanzitutto, chiarito che i problemi etici sono definibili come problemi di scelta tra alternative di comportamento o linee di azione destinate a produrre effetti opposti. Si possono porre problemi di questo tipo nella vita quotidiana. Ad esempio, se ricorrere o meno alla procreazione medicalmente assistita, se rifiutare o meno un trattamento medico salvavita, se donare o meno il sangue, un organo e gli esempi potrebbero moltiplicarsi.

In queste situazioni, chi si confronta con scelte etiche cercherà di trovare una strada per offrire la migliore risposta, basandosi su quel bagaglio di valori

che ha costruito nel suo percorso esistenziale. Tuttavia, le tematiche richiamate a scopo esemplificativo non hanno una rilevanza solo individuale, ma collettiva e come tale istituzionale. Pertanto, occorre capire se quando se ne parla in campo istituzionale e nei media, che fanno spesso da portavoce delle istituzioni, ci stiamo riferendo a problemi di scelta la cui soluzione sta nell'etica diffusa o comune. Come vedremo fra poco non è così. E la differenza che intercorre tra la scelta individuale e quella collettiva sul piano dell'etica pubblica non è di poco conto⁸.

Nel contesto istituzionale, i problemi etici sono problemi di scelta, che possono e devono essere affrontati alla luce di principi e norme, che sono il frutto di una rielaborazione critica dell'etica diffusa. In altre parole, ci si affida all'etica critica. Questa può essere definita come l'ambito dell'agire umano in cui si presentano diversi corsi d'azione tra cui va operata una scelta, per la cui definizione si fa ricorso a criteri che appartengono al mondo del normativo e che richiamano sia valori morali sia criteri giuridici che deontologici⁹.

I principi e le norme giuridiche nel contesto europeo sono espressi dalla cornice dei diritti fondamentali, che considera la dignità umana come valore minimo condiviso e su cui si fonda la dimensione egualitaria di tali diritti.

Diversamente dal contesto istituzionale appena richiamato, dove sono i principi costituzionalizzati della dignità, dell'integrità, della libertà, dell'eguaglianza e della solidarietà a guidare i decisori politici nell'individuazione delle strade da seguire, nella discussione mediatica delle questioni etiche sollevate dall'innovazione scientifica e tecnologica i contorni dei criteri utilizzabili non sono sempre immediati e distinguibili. Questo contesto risente, invero, fortemente dell'uso che si fa nel linguaggio ordinario della nozione di etica, dove è in genere inteso come sinonimo di morale. In tale contesto discorsivo, quando ci si riferisce al concetto di morale vi è spesso implicito il richiamo alla morale corretta, oggettivamente fondata. È raro che nei discorsi comuni *la morale* sia considerata nella sua dimensione storica e culturale. A fronte di tale frequente situazione, quando nella comunicazione mediatica si richiamano problemi etici senza approfondire o ridefinire il termine, risulta piuttosto alto il rischio che il destinatario dell'informazione dia per scontato che si tratti di questioni che trovano risposte in una morale corretta e in una posizione morale oggettiva. Sfugge completamente che le questioni etiche che ci troviamo ad affrontare oggi, siano esse in temi essenzialmente controversi come il fine vita o l'inizio vita, siano esse in tema di IA,

⁸ U. SCHÜKLENK, P. SINGER (a cura di), *Bioethics. An Anthology*, New Jersey, Wiley Blackwell, 2022. Come sostengono gli autori nell'introduzione al volume «[d]iscussing a question at the level of law and public policy [...] raises somewhat different considerations than a discussion of personal ethics, because the consequences of adopting a public policy generally have much wider ramifications than the consequences of a personal choice», p. 7.

⁹ P. BORSELLINO, *Bioetica tra 'moralì' e diritto*, Milano, Raffaello Cortina, 2018.

sono scelte di etica pubblica che possono essere affrontate con un buon uso di criteri normativi, che sono il prodotto culturale e giuridico di una certa epoca storica.

Si tratta di un'epoca storica in cui l'innovazione scientifica e tecnologica ha trasformato radicalmente, migliorandole, le condizioni di vita di molti esseri umani. Tuttavia, nel fare ciò l'innovazione tecnologica ha prodotto un'eccedenza di inedite possibilità di azione che richiedono di essere analizzate con mente aperta. Quello che, invece, attraverso i media spesso si verifica è la presentazione del dibattito sulle questioni etiche in forma polarizzata tra due posizioni estreme. Da una parte, si presentano le posizioni dei bioconservatori, per i quali le questioni etiche sono spesso questioni le cui risposte sono da ricercare nella morale corretta, oggettivamente fondata. La tecnologia è per loro funzionale a mantenere uno status quo e può essere usata solo per mantenere un ordine naturale delle cose che preesiste a tutto. Dall'altra parte, ci sono i transumanisti, che spingono all'estremo l'impiego tecnologico sull'umano, senza dare conto delle ricadute etiche sottese alle loro proposte¹⁰. Questa polarizzazione del dibattito «non dà alcuna vera indicazione sul modo di governare la fase interamente nuova nella quale l'umanità è già entrata»¹¹.

In realtà per governare questa nuova fase occorrono criteri normativi che si sono sviluppati nei diversi ambiti della morale, del diritto e delle etiche settoriali. Le scelte di etica pubblica sono un complesso procedimento che, partendo dal rispetto per la composizione eticamente plurale della società, non deve dimenticare il valore minimo su cui misurare le azioni pubbliche, che è il rispetto per la dignità umana. Questo è sia un valore sia una misura imprescindibile per le scelte pubbliche delle istituzioni europee che vogliono governare la *mutata natura dell'agire umano*¹², superando il vincolo del presente per garantire che lo sviluppo tecnologico rimanga in linea con le esigenze future dell'umanità. Hans Jonas l'ha definita *etica del futuro*¹³, un'etica in grado di ragionare in prospettiva, per tutelare gli interessi delle presenti e future generazioni. Per fare ciò, bisogna innanzitutto sviluppare l'attitudine all'individuazione di questioni etiche, operazione né facile né scontata.

Ogni ricerca di soluzione a un problema etico richiede di tenere presente che spesso, anche se non sempre, i problemi etici si nascondono dietro problemi tecnici o dipendono dall'esistenza di problemi tecnici. Spesso chi detiene la conoscenza esperta su un dato tema, ad esempio gli informatici, fanno appello al

¹⁰ Per un approfondimento sul movimento transumanista cfr. F.H. LLANO ALONSO, *Homo excelsior. Los límites ético-jurídicos del transhumanismo*, Valencia, Tirando Lo Blanch, 2018; J. TESTART, A. ROUSSEAU, *Au péril de l'humain*, Paris, Éditions du Seuil, 2018.

¹¹ S. RODOTÀ, *Il diritto di avere diritti*, Roma-Napoli, Laterza, 2012, p. 355.

¹² H. JONAS, *Il principio di responsabilità*, cit. p. 16 ss.

¹³ *Ibidem*.

loro mandato tecnico-scientifico e alla neutralità di ciò che fanno e progettano per dichiararsi estranei alle questioni etiche e a un eventuale dovere di affrontarle. Questo atteggiamento è comune agli esperti di diversi saperi tecnici e scientifici. In bioetica, questo fenomeno è stato studiato ed è stato opportunamente definito con l'espressione il *paradosso di un'etica che dissimula l'etica*¹⁴. Tale paradosso si è reso evidente, già da diverso tempo, nei contesti di studio più diversi: dalle tematiche tradizionali della riflessione bioetica¹⁵ a quelle ambientali¹⁶. Dietro la concezione del mandato esclusivamente tecnico¹⁷ di certe professioni e discipline si cela spesso una soluzione etica implicita circa la spettanza delle decisioni e degli obiettivi da perseguire così come il risultato atteso. Detto diversamente, nell'attuale mondo tecnologico, vi è un assunto, quasi mai esplicitato, secondo il quale la soluzione tecnica è *la* soluzione per risolvere i nostri problemi ed è quella che va, sempre e comunque, perseguita. Come ci ricorda Garrett Hardin, però, questo è un atteggiamento che caratterizza la società umana da un preciso momento storico, ossia da quando la tecnologia ha assunto un ruolo cruciale nello sviluppo economico e sociale¹⁸. Tuttavia, come Hardin stesso avrà modo di sottolineare, tali soluzioni non sono necessariamente sempre *la* soluzione a lungo termine di un problema.

Hardin definisce il problema tecnico come un problema che «requires a change only in the techniques of the natural sciences, demanding little or nothing in the way of change in human values or ideas of morality»¹⁹. Di fatto, si ha

¹⁴ P. BORSELLINO, *Bioetica tra 'moralì' e diritto*, cit.

¹⁵ Ricordiamo che lo sviluppo tecnologico degli anni '60-'70 del secolo scorso nel contesto statunitense ha fatto emergere i problemi etici relativi al fine vita. In particolare, a fronte della disponibilità di tecnologie salvavita che consentono di prolungare, anche per un tempo molto lungo, il processo del morire (fattibilità tecnica), ci si chiese se fosse sempre lecito il loro impiego indipendentemente dalla volontà del paziente (sostenibilità etica) e se unico decisore dovesse restare il clinico.

¹⁶ Garrett Hardin fu uno dei primi a sottolineare che, in relazione ai beni comuni, la soluzione non è tanto di tipo tecnico, quanto di tipo etico. Cfr. G. HARDIN, *The Tragedy of the Commons*, in *Science*, 1968, 162, 3859, pp. 1243-1248.

¹⁷ La riflessione etica sulla relazione di cura nelle situazioni di fine vita, iniziata negli anni '70 del secolo scorso, ha messo in evidenza come, per lungo tempo, i clinici abbiano pensato di svolgere un'attività tecnico-scientifica, non implicante questioni etiche e scelte morali. Tuttavia, proprio dietro questa concezione della professione del medico si nascondevano precisi assunti etici come di seguito elencati: 1) che solo al medico spettino le decisioni in quanto detentore del sapere e delle competenze tecnico-scientifiche per decidere ed è quindi sempre in grado di prendere decisioni adeguate per il bene del paziente; 2) che gli interventi terapeutici abbiano un solo obiettivo, quello di prolungare la vita dei pazienti, a tutti i costi; 3) che il medico occupi una posizione di supremazia nella relazione di cura che gli permette di decidere se e quando mettere al corrente il paziente delle scelte terapeutiche. Cfr. P. BORSELLINO, *Bioetica tra 'moralì' e diritto*, cit.

¹⁸ G. HARDIN, *The Tragedy of the Commons*, cit. p. 1243. Come osserva l'autore «[i]n our day (though not in earlier times) technical solutions are always welcome».

¹⁹ *Ibidem*.

un problema tecnico quando per la sua risoluzione si può ricorrere a competenze specialistiche o a strategie istituzionali, ad esempio, di tipo organizzativo o a conoscenze ancora in divenire, sulla base delle quali si può ragionevolmente prevedere che, a breve, una data questione tecnica sarà risolta²⁰.

Tuttavia, la storia dello sviluppo tecnologico ha mostrato come la ricerca di una soluzione *sempre e solo* sul piano tecnico risulti limitativa rispetto al pluriverso dimensionale dell'esistenza umana.

In sintesi, se per risolvere un problema, ci si arresta alla ricerca della sola soluzione tecnica si rischia di non vedere che, spesso, dietro al problema tecnico si celano precisi assunti etici che riguardano il potere decisionale e la scelta precostituita della direzione in cui procedere. Alcuni degli assunti etici dissimulati dietro la soluzione tecnica, per ciò che riguarda l'oggetto di interesse di questo lavoro, sono i seguenti: 1) la tecnologia non è né buona né cattiva; 2) essa evolve naturalmente per lo scopo per cui è creata, ossia essere usata da chiunque; 3) ha in sé una vocazione egalitaria; 4) non ha bisogno di regole etico-giuridiche, basta l'autoregolamentazione perché il fine è il benessere dell'umanità; 5) lo sviluppo tecnologico è inevitabile, quindi il destino dell'umanità è nelle mani della tecnologia; 6) i dati in generale, e in particolare i nostri, servono alla tecnologia per migliorare sé stessa in funzione dei servizi che ci offre²¹.

Una significativa ricaduta negativa dell'approccio, prevalentemente, se non esclusivamente, tecnico-scientifico all'IA è di supportare l'idea che l'IA sia eticamente neutrale. In realtà, lungi dall'essere neutrali, i sistemi di IA sono programmati per compiere scelte a cui sono sottesi precisi valori. È questo il caso, ad esempio, delle automobili a guida autonoma²². Esse vengono programmate sulla base di regole necessarie per operare scelte in situazioni moralmente impegnative. Tuttavia, è la mano del programmatore che decide quali saranno le regole inserite e, in ultima istanza, se le regole saranno etiche o meno. In altre parole, mentre tali sistemi possono essere programmati inserendo regole etiche, essi non comprendono la qualità etica di quelle regole e come è stato affermato «data-based systems would respect non-ethical or unethical rules in the same manner»²³.

²⁰ Si pensi, a titolo esemplificativo, a cosa ha significato l'avvento del *Deep Learning* nell'apprendimento automatico senza pre-processamento dei dati.

²¹ Avremo modo di esemplificare quanto sinora sostenuto nel capitolo dedicato ai casi di studio dell'IA in medicina e dell'IA nel dibattito sul potenziamento umano.

²² Cfr. S. SALARDI, *Autonomia e responsabilità alla prova di autonomus artificial systems. Profili filosofici, etici e giuridici. Il caso delle driverless cars*, in D. CERINI, A. PISANI TEDESCO (a cura di), *Smart mobility, smart cars e intelligenza artificiale: responsabilità e prospettive*, Torino, Giappichelli, 2019, pp. 117-139.

²³ P. KIRCHSCHLAEGGER, *Digital Transformation and Ethics*, Baden-Baden, Nomos, 2021, p.

A muovere dalle considerazioni che precedono, possiamo affermare che se si è convinti che le questioni tecniche siano risolvibili sempre e solo con strumenti tecnici, è giocoforza dedurre che chi può decidere è solo chi possiede le competenze tecniche. Inoltre, se si sposa l'idea che le soluzioni utili a risolvere problemi e ad affrontare sfide come, ad esempio, quelle ambientali sono solo quelle tecniche, allora non vi sarà necessità di individuare strade alternative, perché la strada obbligata sarà quella, e solo quella, dell'innovazione nella limitata declinazione tecnologica. In questo scenario non entra nemmeno in discussione l'ipotesi che, in certi settori di attività umana, possa essere meglio procedere lentamente e con proposte alternative al *mainstream*, mantenendo lo *status quo* sul piano tecnologico, considerati i fini e gli obiettivi perseguiti²⁴, come auspica, ad esempio, Frank Pasquale²⁵.

In pratica, quando la ricerca di soluzioni a problemi sollevati dalle attività umane, dalla loro organizzazione e struttura, nonché dalla loro pianificazione futura, si ferma alla soluzione tecnica, appare scontato individuare nei tecnici i soggetti legittimati a prendere decisioni, anche quando queste non sono di carattere tecnico, bensì politico ed etico. Coloro che ritengono che i problemi, che l'umanità deve affrontare oggigiorno – cambiamento climatico, impoverimento generalizzato, globalizzazione ecc. –, siano solo problemi tecnici da risolvere con sapere tecnico sposano l'idea che bastano nuove tecnologie, metodi e procedure tecniche e che l'avanzamento scientifico-tecnologico troverà sempre una soluzione adeguata alle problematiche che, di volta in volta, si presentano. In questa prospettiva, non si dà conto del fatto che spesso la soluzione approntata grazie all'evoluzione delle tecniche difficilmente è totalmente risolutiva, in quanto porta con sé nuovi e inaspettati problemi etici, di norma dovuti alla maggiore capacità delle tecnologie più recenti di incidere in modo pervasivo sulla vita umana e non umana²⁶.

In poche parole, chi si nasconde dietro l'esclusivo riferimento alle soluzioni tecniche sta operando un *assalto alla nostra consapevolezza*, per dirla con Zuboff²⁷

113. Sul tema della cybersicurezza vedasi gli articoli di C. GIUSTOZZI reperibili a questo link <https://www.agendadigitale.eu/giornalista/corrado-giustozzi/>.

²⁴ Nel contesto educativo, ad esempio, si sostiene che «[...] technological applications should only be introduced or used in education if they make sense from an ethical, from a pedagogical, and from a didactic perspective – not just because they represent innovations or because they are sponsored by, or are a gift from, the private sector», cfr. P. KIRCHSCHLAEGER, *Digital Transformation and Ethics*, cit. p. 391.

²⁵ F. PASQUALE, *Le nuove leggi della robotica. Difendere la competenza umana nell'era dell'intelligenza artificiale*, Roma, LUISS University Press, 2020.

²⁶ Molti possono essere gli esempi, ci limitiamo a ricordare che in campo genetico questo è avvenuto con l'avvento della tecnica CRISPR-CAS 9.

²⁷ S. ZUBOFF, *Il capitalismo della sorveglianza. Il futuro dell'umanità nell'era dei nuovi poteri*, Roma, LUISS University Press, 2019, p. 324.

e, al contempo, incentiva la diffusione e il radicamento dell'ideologia deterministica della tecnica. Determinismo tecnologico e neutralità della tecnica sono le parole chiave della visione tecnocentrica della società.

Tale visione è supportata dal valore che oggi viene dato ai dati nella ricostruzione del reale. È a partire dal Novecento che il concetto di dato ha perso quella caratteristica di irriducibilità che si poneva come limite alla conoscenza e alla sua validità. Nelle conclusioni di autori come Sellars e Goodman²⁸ non esistono dati che non siano inclusi e derivabili da una sovrastruttura teorica e, pertanto, non esistono dati ultimi su cui fondare, inequivocabilmente, l'attività conoscitiva.

La disponibilità di una quantità mai vista prima di dati, che stimola lo sviluppo di nuove tecnologie, infonde fiducia circa la possibilità di catturare il reale nella sua completezza. Di fatto, è possibile, grazie alla *Data Analytics*, estrarre correlazioni probabilistiche dai dati su cui basare previsioni e, in alcuni casi, prendere decisioni. Dalla nozione di dato si è passati alla nozione di *big data* per indicare «things one can do at a large scale that cannot be done at a smaller one, to extract new insights or create new forms of value, in ways that change markets, organizations, the relationship between citizens and governments, and more»²⁹. I *big data* consentono di fare tutto ciò e di inserirsi in modo pervasivo nella vita sociale e individuale, poiché a monte vi è l'assunto che tutto sia rappresentabile e risolvibile mediante una soluzione tecnica. Seguendo la concezione tecnocentrica si è affermata l'idea che, dato l'enorme volume dei dati generati, essi possano considerarsi rappresentativi di tutta la realtà sociale e naturale, enfatizzando una *rhetoric of exactitude*. Come è stato osservato «[a]ssuming that data are neutral and fully exhaustive, the problem in handling them becomes technical»³⁰. In realtà, questa retorica nasconde questioni etiche profonde. In poche parole, esistono problemi etici dissimulati da problemi tecnici relativi anche ai *big data*³¹. Le quattro V che li caratterizzano, ossia volume, velocità, veridicità (qualità e affidabilità dei dati) e varietà sono esse stesse fonti di criticità etiche. Ad esempio, il volume è in realtà una questione di percezione³², una qualità *storica-*

²⁸Cfr. W. SELLARS, *Empirismo e filosofia della mente*, Torino, Einaudi, 2004 (versione originale 1956); N. GOODMAN, *I linguaggi dell'arte*, Milano, Il Saggiatore, 2017 (versione originale 1978).

²⁹V. MAYER-SCHONBERGER, K. CUKIER, *Big Data: A Revolution That Will Transform How We Live, Work, and Think*, Boston, Harcourt, 2013, p. 10.

³⁰D. BALAZKA, D. RODIGHIERO, *Big Data and the Little Big Bang: An Epistemological (R)evolution*, in *Frontiers in Big Data*, 2020, 3, 31, pp. 1-10, ivi, p. 5.

³¹Vedasi l'approfondimento di Dreyfus sulla c.d. *ontological assumption*, H.L. DREYFUS, *What Computers Still Can't Do: A Critique of Artificial Reason*, Cambridge, The MIT Press, 1992 (versione originale 1972) p. 206. Sul tema della neutralità dei dati cfr. Anche A.C. AMATO MANGIAMELLI, *Algoritmi e big data*, in *Rivista di Filosofia del Diritto*, 2019, VII, 1, pp. 107-124.

³²D. BALAZKA, D. RODIGHIERO, *Big Data and the Little Big Bang: An Epistemological (R)evolution*, cit. p. 2.

mente contestualizzata, dipendente dalle tecnologie disponibili in un dato momento³³. A sua volta, la veridicità dei dati non può essere assunta come innata nei dati stessi, in quanto questi ultimi sono manipolabili da chi li inserisce nei sistemi e possono ereditare *bias* che risalgono al modo in cui sono stati archiviati in passato (incompletezza, limitatezza del campione, ideologia sottesa a precedenti archiviazioni, ad esempio, per campioni etnici, razziali, religiosi ecc.). Si pone, pertanto, un problema di trasparenza nella raccolta, nella gestione e nel controllo degli esiti ottenuti sulla base dei dati con cui si alimenta un sistema di IA. La trasparenza è una questione tanto più urgente da affrontare, quanto più gli scopi che si perseguono hanno ricadute su vasta scala.

Non si è confrontati, quindi, esclusivamente con un problema tecnico, ossia trovare il modo pratico di garantire trasparenza. In realtà, il problema è etico e riguarda la determinazione degli scopi relativi all'uso che si fa dei dati in un certo contesto, che è ciò che determina il senso di quella trasparenza.

Il tema della trasparenza e della correlata necessità di *explainability*³⁴ delle tecnologie intelligenti è stato sollevato in relazione agli algoritmi c.d. *black box*. Sono algoritmi che presentano almeno tre livelli di opacità.

Il primo riguarda l'intelligibilità da parte dell'utente che è destinatario di decisioni automatizzate. A fronte della complessità delle procedure algoritmiche, ci si chiede se e fino a che punto l'utente finale debba comprenderne i meccanismi per potersi difendere da eventuali decisioni ingiuste o inique che lo riguardano³⁵. Il secondo livello concerne l'opacità intenzionalmente indotta nei sistemi di IA da parte di istituzioni, organizzazioni, enti ecc. per ragioni sia lecite che illecite³⁶. In merito a tale livello, ci si chiede quale possano essere le situazioni in cui questa opacità possa essere legittimata.

Infine, il terzo livello si riferisce a quegli algoritmi che, dato un certo input, arrivano a un risultato finale (output) di cui non è possibile ripercorrere chiaramente i passaggi che lo hanno prodotto³⁷. In questo caso, nemmeno il programmatore è in grado di comprendere esattamente come si sia giunti all'esito finale.

³³ *Ibidem*.

³⁴ Sulle caratteristiche di questi algoritmi cfr. l'articolo intitolato *Opening the 'black box' of artificial intelligence*, dicembre 2020 accessibile al link <https://ec.europa.eu/research-and-innovation/en/horizon-magazine/opening-black-box-artificial-intelligence>.

³⁵ Ad esempio, le auto autonome possiedono sensori utili a permettere loro di guidare. Tali sensori, tuttavia, possono presentare delle carenze, che se non comprese, possono indurre, nell'utente finale, errori di scelta circa quando utilizzare la guida autonoma.

³⁶ Ad esempio, un'agenzia di intelligence usa sistemi intelligenti per individuare soggetti pericolosi in contesti pubblici, ma si rifiuta di rendere noti i criteri di classificazione usati dall'algoritmo impiegato.

³⁷ Come chiarisce Massimo Chiriatti «da un punto di vista matematico, il deep learning è del tutto trasparente: consiste in una serie di moltiplicazioni di matrici a cui si applicano funzioni non

Mentre per i primi due livelli, il problema della trasparenza si gioca, innanzitutto, attorno a scelte di convenienza o di corretto bilanciamento di valori, l'ultimo livello di opacità è quello più insidioso perché un algoritmo indecifrabile nel suo processo decisionale diviene un importante ostacolo per elaborare strategie di trasparenza anche nel caso di opacità di primo e secondo livello.

Una possibile soluzione tecnica ai problemi di trasparenza degli algoritmi *black box* è quella di adottare algoritmi c.d. *white box*³⁸. Questi ultimi sono algoritmi la cui logica interna e i passaggi del processo decisionale sono chiari e intelleggibili dalla progettazione, come è il caso dei c.d. alberi decisionali (*decision tree*).

Sebbene questi algoritmi presentino questo grande vantaggio, essi hanno altresì lo svantaggio di poter essere usati per attività predittive meno complesse.

In contesti in cui l'attività predittiva è complessa devono essere usati gli algoritmi *black box*, che non sono decifrabili sin dalla programmazione, ma devono essere resi intelleggibili attraverso l'applicazione di tecniche per estrarre informazioni dalla logica interna o dai risultati del modello. Il processo che porta alla trasparenza richiede, pertanto, tempo e accuratezza nell'analisi. Nel frattempo, però, l'impiego di questi algoritmi, in cui possono essere presenti *bias* originari, continua e solleva importanti questioni etiche.

Nell'ambito della giustizia, ad esempio, il problema etico, che si cela dietro l'opacità di alcuni algoritmi, può incidere su principi e valori costituzionalmente garantiti come l'eguaglianza e il divieto di discriminazione³⁹. Ma oltre a questa

lineari. Tuttavia, il suo risultato finale è probabilistico, nel senso che non si può essere certi che il risultato sia corretto», M. CHIRIATTI, *Incoscienza artificiale. Come fanno le macchine a prevedere per noi*, cit. p. 59.

³⁸PANEL FOR THE FUTURE OF SCIENCE AND TECHNOLOGY (STOA), *Understanding Algorithmic Decision-Making: Opportunities and Challenges*, 2019, accessibile al seguente link [https://www.europarl.europa.eu/RegData/etudes/STUD/2019/624261/EPRS_STU\(2019\)624261_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2019/624261/EPRS_STU(2019)624261_EN.pdf)

³⁹È noto il caso dell'algoritmo COMPAS, usato negli USA per determinare il rischio di recidiva di un imputato in un processo penale. Nel 2016 la Corte Suprema del Wisconsin si è pronunciata nella vicenda del sig.re Loomis. Il soggetto, trovato alla guida di un'auto usata per una sparatoria, fu condannato a una pena molto severa sulla base della valutazione operata da COMPAS sul rischio di recidiva e pericolosità per la comunità. La difesa aveva sollevato la questione dell'accuratezza della valutazione di COMPAS, in quanto limitato ad alcuni parametri come l'etnia, le condizioni sociali, il livello di educazione e il background familiare.

Cfr. State Vs. Loomis, 881 N.W. 2d 749 (Wis. 2016). Sul caso Loomis cfr. S. CARRER, *Se l'amicus curiae è un algoritmo: il chiacchierato caso Loomis alla Corte Suprema del Wisconsin*, in *Giurisprudenza Penale Web*, 2019, 4, accessibile al link <https://www.giurisprudenzapenale.com/2019/04/24/lamicus-curiae-un-algoritmo-chiacchierato-caso-loomis-alla-corte-suprema-del-wisconsin/>.

Ma ovviamente vi sono ulteriori sfide nell'era digitale, cfr. S. VANTIN, *Il diritto antidiscriminatorio nell'era digitale*, Milano, Wolters Kluwer, 2021; T. CASADEI, S. PIETROPAOLI (a cura di), *Diritto e Tecnologie Informatiche*, Milano, Wolters Kluwer, 2021; F. FAINI, S. PIETROPAOLI, *Scienze giuridica e tecnologie informatiche*, Torino, Giappichelli, 2021; C. FARALLI (a cura di), *Forum Vulnerabilità e Nuove Tecnologie*, in *Notizie di Politeia*, 2019, 136, pp. 5-108.

Sul caso in questione cfr. anche A. SIMONCINI, S. SUWEIS, *Il cambio di paradigma nell'intelli-*

criticità, l'opacità di alcuni algoritmi rappresenta un grosso limite procedurale perché può rendere difficoltosa la difesa in giudizio. Infatti, un algoritmo opaco determina la preventiva impossibilità di spiegare in giudizio il ragionamento con cui si è arrivati a una certa conclusione, rendendo così vano il tentativo di spiegare lo scopo dell'algoritmo al giudice.

Anche in campo medico, diagnosi, basate su risultati di algoritmi *black-box*, possono incidere sulla qualità e la sostanza dell'informazione da parte del medico al paziente, con conseguenze importanti sull'autonomia delle scelte del paziente. Dedicheremo un approfondimento a questo tema nel capitolo quarto.

In sintesi, i problemi tecnici possono nascondere problemi etici, la cui risoluzione richiede la messa in campo di criteri normativi e nuove categorie come la trasparenza, l'esplicabilità (*explicability*) e l'affidabilità (*trustworthiness*).

Per completezza occorre concludere le riflessioni di questo paragrafo richiamando brevemente il c.d. *dataism*, ossia una nuova filosofia o addirittura religione⁴⁰, che riduce tutta l'esperienza umana e naturale ai dati. In questa prospettiva, i dati rappresentano tutta la realtà -umana e naturale- e ciò che risulta dai dati è certo ed esatto.

Come vedremo nel prossimo paragrafo, la diffusione di una siffatta ideologia (*data-ideology*) è possibile per molti motivi. Ma ve ne è uno che gioca un ruolo centrale al di fuori dei contesti specializzati. Si tratta della non chiarita differenza tra conoscenza attraverso il metodo scientifico e conoscenza attraverso le analisi dei dati da parte di sistemi intelligenti. In altre parole, il non esplicitato passaggio dalla causalità alla correlazione. Di ciò parleremo nel seguente paragrafo.

4. La conoscenza nella prospettiva tecnologica e in quella scientifica

La conoscenza dell'esperienza possibile, unitamente alla ricerca di una spiegazione di cosa sia il reale e di quale sia la sua natura oltre l'apparenza, sono elementi costitutivi della questione della conoscenza, uno dei più antichi temi nella storia del pensiero filosofico. In poche parole, ci si è da sempre chiesti se esista una realtà empirica oggettiva, passibile di descrizione pura, ossia non *deformata* dalla soggettività di chi la osserva⁴¹.

genza artificiale e il suo impatto sul diritto costituzionale, in *Rivista di Filosofia del Diritto*, 2019, VII, 1, pp. 87-106.

⁴⁰ Il termine fu coniato per la prima volta da David BROOKS nel suo articolo intitolato *The Philosophy of Data*, apparso nel 2013 sul *NY Times*, <https://www.nytimes.com/2013/02/05/opinion/brooks-the-philosophy-of-data.html>. Viene ripreso e reinterpretato da Yuval Noah HARARI nel 2017 nell'articolo intitolato *Dataism is Our New God*, in *New Perspectives Quarterly*, 34, 2, pp. 36-43.

⁴¹ Si era posto questo problema già Immanuel Kant, il quale scrive «[q]uando una cosa è vali-

Intendiamo, in via generale, per conoscenza «una tecnica per l'accertamento di un oggetto qualsiasi, o la disponibilità o il possesso di una tecnica siffatta. Per tecnica di accertamento va intesa una qualsiasi procedura che renda possibile la descrizione, il calcolo o la previsione controllabile di un oggetto; e per oggetto va intesa qualsiasi entità, fatto, cosa, realtà o proprietà, che possa essere sottoposto a una tale procedura»⁴². Il rapporto tra l'oggetto, di volta in volta, selezionato e l'operazione conoscitiva che lo riguarda è stato interpretato in maniera differente nel corso della storia della filosofia⁴³. I termini protagonisti di questa storia sono *trascendente* e *trascendentale*. Tagliando molto attraverso la loro storia semantica, possiamo dire che fino a Kant i due termini indicano «le proprietà che tutte le cose hanno in comune, e che perciò eccedono o trascendono la diversità dei generi in cui le cose si distribuiscono»⁴⁴. A partire da Kant, *trascendente* verrà a indicare ciò che eccede i limiti dell'esperienza possibile⁴⁵, mentre *trascendentale*, verrà a significare «ciò che antecede l'esperienza (a priori) pur non essendo destinato ad altro che a rendere possibile la semplice conoscenza empirica»⁴⁶.

La teoria della scienza c.d. metafisica si fonda sulla concezione essenzialistica e trascendente della realtà.

Nel capitolo precedente, abbiamo affrontato la teoria essenzialistica con riferimento al linguaggio, qui la riprendiamo in un diverso contesto, in quanto è stata una delle fonti che ha guidato a lungo la conoscenza scientifica. Secondo tale concezione, vi sarebbero verità ultime che lo scienziato è in grado di stabilire definitivamente e le «teorie migliori, quelle veramente scientifiche, descrivono l'«essenza», o la «natura essenziale» delle cose: le realtà che giacciono al di là delle apparenze»⁴⁷. La metafisica è, quindi, la scienza *prima* che antecede tutte le altre scienze e ne determina la validità. Obiettivo della metafisica è, invero, la spiegazione ultima e incontrovertibile, che si fonda su una tendenza aprioristica o speculativa. L'approccio metafisico, che assume delle verità fisse da scoprire,

da per qualunque essere in possesso di ragione, la sua ragion d'essere è oggettiva e sufficiente», cfr. I. KANT, *Critica della Ragion Pura*, Lipsia, Hartnack Riga, 1781, p. 848.

⁴² N. ABBAGNANO, *Dizionario di Filosofia*, Torino, Utet, 2013, p. 193.

⁴³ Ci ricorda Mario Calderoni che «[d]ue furono in ogni tempo le vie per cui l'uomo tentò di accrescere la sfera della propria conoscenza: ora ripiegandosi sopra se stesso, traendo dalle profondità della propria mente e dalla contemplazione delle idee la parte maggiore dello scibile; ora invece volgendo lo sguardo attento allo svolgersi dei fatti nel mondo reale, registrandoli con pazienti osservazioni e confronti, per ridurli a formule via via più schematiche e generali», M. CALDERONI, *I postulati della scienza positiva ed il diritto penale*, in G. PAPINI (a cura di), *Scritti di Mario Calderoni*, vol. 1, Società Anon Editrice 'La Voce', Firenze, 1924, p. 34.

⁴⁴ N. ABBAGNANO, *Dizionario di Filosofia*, cit. p. 1117.

⁴⁵ *Ibidem*, p. 1119.

⁴⁶ *Ibidem*.

⁴⁷ K. POPPER, *Scienza e Filosofia*, Torino, Einaudi, 1969, p. 21.

porta con sé un inevitabile atteggiamento dogmatico. Questo atteggiamento può essere spiegato considerando che, in mancanza di un metodo di replicazione e riconferma dei risultati raggiunti, come reso possibile dal principio di falsificazione introdotto da Popper, la cristallizzazione dei risultati ottenuti appare un modo «per proteggersi e per perpetuare valori raggiunti»⁴⁸. In pratica, fino ad un certo punto della storia della scienza, assumere verità prime su cui fondare il progresso della conoscenza serve da garanzia di certezza, in assenza di metodi di verifiche e controlli⁴⁹, che sarebbero stati elaborati successivamente⁵⁰.

La concezione essenzialistica fu parte costitutiva della filosofia galileiana⁵¹, che Popper si rifiutò di difendere. Per Popper, che ha portato a compimento la critica alla metafisica, invece «[L]e teorie sono reti gettate per catturare quello che noi chiamiamo “il mondo”: per razionalizzarlo, per spiegarlo, per dominarlo»⁵².

Sarà con l'apporto teorico di Karl Popper che la conoscenza scientifica assumerà le caratteristiche empirico-sperimentali che la definiscono attualmente, smarcandosi definitivamente dalle influenze metafisiche.

Il criterio individuato dall'autore per distinguere la conoscenza scientifica (criterio di demarcazione) dalla metafisica è il noto procedimento di falsificazione. Le teorie scientifiche sono falsificabili, mentre quelle metafisiche non lo sono⁵³. Il criterio proposto da Popper sostituisce quello di verificabilità di origine neopositivista. Secondo il principio di falsificabilità, una teoria è empirica se può essere smentita dall'esperienza e tale principio presenta una “superiorità” epistemologica rispetto a quello di verificabilità, perché miliardi di conferme sul piano empirico non possono rendere certa una teoria, mentre basta una smentita per confutarla irrimediabilmente.

Il superamento della metafisica non è un problema solo metodologico.

In senso più generale, la linea di confine tra la cultura moderna e quella premoderna è la critica radicale della visione metafisica propria della prospetti-

⁴⁸J. DEWEY, *Alcuni stadi del pensiero logico*, in R. FREGA (a cura di), *Logica Sperimentale, Teoria Naturalistica della Conoscenza e del Pensiero*, Macerata, Quodlibet, 2008, pp. 3-22, ivi, p. 7.

⁴⁹*Ibidem*.

⁵⁰K. POPPER, *Logica della scoperta scientifica*, Torino, Einaudi, 1998 (prima edizione originale 1934), p. 39. Secondo l'autore «[l']essenzialismo guarda al nostro mondo ordinario come a una pura e semplice apparenza, dietro la quale esso – l'essenzialismo – scopre il mondo reale. Questa teoria dev'essere messa da parte non appena siamo diventati consapevoli del fatto che il mondo di ciascuna delle nostre teorie può essere spiegato, a sua volta, da mondi ulteriori, descritti da ulteriori teorie: da teorie situate a un livello più alto di astrazione, di universalità e di controllabilità. La dottrina di una *realtà essenziale o ultima* crolla insieme con quella di una spiegazione ultima».

⁵¹K. POPPER, *Scienza e Filosofia*, cit. p. 21.

⁵²K. POPPER, *Logica della scoperta scientifica*, cit. p. 43.

⁵³K. POPPER, *Logica della scoperta scientifica*, cit.; M. MARSONET, *I rapporti tra scienza e metafisica*, in *Acta Philosophica*, 1996, 5, 2, pp. 251-268.

va medievale cristiana ed aristotelica. Si critica l'idea che tutto abbia necessariamente un ordine, un posto e un fine predeterminati nella natura come nella società. In questa prospettiva «la distribuzione dei compiti manifesta dunque, nell'aspetto soggettivo, la natura e il destino dei vari esseri umani; mentre, per l'aspetto oggettivo, il distinguersi delle conoscenze e delle tecniche dovrà rispecchiare l'ordine della realtà e le essenze delle cose, colte mediante un sistema di definizioni reali»⁵⁴. Il punto di frattura tra questa prospettiva e quella propriamente scientifica nell'accezione moderna è che l'indagine si colloca nel perimetro dell'esperienza e non lo travalica alla ricerca di una struttura trascendente. Il metodo empirico-sperimentale, che si avvale di una modalità di ragionamento deduttivo⁵⁵, è quello atto a meglio garantire la falsificabilità empirica e, quindi, il controllo intersoggettivo delle teorie scientifiche. Questo tipo di ragionamento differisce da quello applicato dagli scienziati del Seicento.

La letteratura scientifica di quell'epoca fonda la conoscenza sull'osservazione che, per mezzo del metodo induttivo, consente di arrivare a leggi generalizzabili. Come è stato osservato «[a] partire da Bacone, infatti, tutta l'attenzione viene rivolta all'osservazione, ai fatti, al coinvolgimento con il mondo esterno»⁵⁶. La costruzione del metodo empirico-sperimentale, basato sul ragionamento deduttivo, è una realtà relativamente recente.

Ci sono voluti secoli e passaggi intermedi di specificazione da parte di importanti pensatori per arrivare a superare il metodo induttivo. Ad esempio, René Descartes⁵⁷ propone l'intuizione come passaggio per arrivare alla formulazione di un'ipotesi; poi Claude Bernard⁵⁸ pone il problema della distinzione tra avere un'esperienza e condurre un esperimento. Sarà questo il discrimine tra ciò che consideriamo metodo scientifico, ossia quel metodo fondato sul rigoroso controllo intersoggettivo, che consente di andare alla ricerca di fatti che potrebbero smentirla, e ciò che sta fuori da questa definizione. Seguendo tale prospettiva, «[l]a conoscenza diventa un processo relativo, in cui non esistono dati assoluti, ma solo un progressivo modificarsi»⁵⁹. Questa svolta ha agevolato il passaggio

⁵⁴U. SCARPELLI, *Scienza, Sapere, Sapienza*, in *Rivista Internazionale di Filosofia del Diritto*, 1986, IV Serie, LXIII, pp. 245-258, p. 246.

⁵⁵Occorre ricordare che al di là dell'applicazione nel campo scientifico, esiste una capacità deduttiva umana che si fonda su un meccanismo rapido, basato sull'impiego di conoscenze generali e che mira a «raggiungere conclusioni vere, o almeno plausibili, e non semplicemente fare deduzioni valide». Cfr. P.N. JOHNSON-LAIRD, *Deduzione, Induzione, Creatività. Pensiero Umano e Pensiero Meccanico*, Bologna, Il Mulino, 1994, p. 37.

⁵⁶J. DEWEY, *Alcuni stadi del pensiero logico*, cit. p. 10.

⁵⁷R. DESCARTES, *Discorso sul metodo*, Brescia, Editrice la Scuola, 1983.

⁵⁸C. BERNARD, *Introduzione allo studio della medicina sperimentale*, Padova, Piccin, 1994.

⁵⁹L.R. ANGELETTI, V. GAZZANIGA, *Storia, filosofia ed etica generale della medicina*, Milano, Elsevier, 2012, p. 170.

da una visione realista a una visione costruttivista della conoscenza, che consente di mettere in dubbio pretese assolutiste circa l'oggettività del progresso scientifico rispetto al contesto storico e socioculturale in cui avviene. Come è stato affermato «[s]cientific knowledge, it is now widely accepted, does not simply accumulate, nor does technology invariably advance benign human interests. Changes in both happen within social parameters that have already been laid down, often long in advance»⁶⁰. Porre l'accento sulla visione costruttivista consente di fare emergere un presupposto, spesso nascosto, circa il rapporto tra scienza e *Verità*. Di fatto, possiamo certamente affermare che l'attività conoscitiva della scienza consente la ricerca della verità. Se ciò è vero, è altresì vero che, se questa ricerca non è intesa come semplice accumulo di conoscenze, ma come scelta di indagare certi ambiti dell'esperienza scartandone altri, la verità raggiunta è circoscritta ai soli ambiti di indagine e, quindi, non è assoluta. Essa è riferita al contesto di indagine e rispetto a quello è una verità controllabile in quanto ripetibile e falsificabile.

Ciò perché il metodo scientifico si basa su ipotesi corroborabili, elaborate dagli scienziati, che cercano di capire le ragioni e i meccanismi di collegamento tra un evento e una causa⁶¹. Lo studio della *causalità* è stato al centro dell'attenzione filosofica già a partire da Platone⁶² e ha conosciuto nel corso della storia del pensiero filosofico diverse declinazioni. Nell'accezione deterministica, *strict determinism*, il principio di causalità consiste in una relazione necessaria tra causa ed evento. Questo significato di stretta causalità è stato superato nello stesso ambito scientifico con il *principio di indeterminazione* di Heisenberg⁶³. Tale principio ha obbligato ad abbandonare la concezione della causalità come necessità.

⁶⁰ S. JASANOFF, *Designs on Nature*, Princeton and Oxford, Princeton University Press, 2005, p. 13. Le suggestioni metafisiche si ripropongono oggi in contesti postmoderni e si pongono come critica al costruttivismo sociale tipico di quest'epoca. Nella seconda metà degli anni Novanta del 1900 la *Cybernetic Culture Research Unit* dell'Università di Warwick si muove in questa direzione. Tra i promotori di un rinnovato approccio metafisico alla realtà vi sono Sadie Plant e Nick Land.

⁶¹ Karl Popper sostiene chiaramente, alla luce degli sviluppi in fisica del tempo (principio di indeterminazione di Heisenberg), di non adottare né respingere il principio di causalità. Egli adotta un principio ad esso molto vicino, ossia che «non dobbiamo abbandonare la ricerca di leggi universali e di sistemi coerenti di teorie, né dobbiamo rinunciare ai nostri tentativi di spiegare causalmente qualunque tipo di evento che siamo in grado di descrivere», K. POPPER, *Logica della scoperta scientifica*, cit. pp. 45-46.

⁶² Platone «considera la causa come il principio per il quale una cosa è, o diventa, ciò che è», cfr. N. ABBAGNANO, *Dizionario di Filosofia*, cit. p. 143.

⁶³ Come noto «[t]ale principio dice che vi è un limite alla precisione con cui si può studiare il moto di un oggetto piccolissimo» cfr. R.E. PEIERLS, *Le leggi della natura*, Torino, Boringhieri, 1960, p. 202. La dimostrazione che le leggi fisiche fondamentali non hanno un fondamento de-

In ambito filosofico⁶⁴, invece, il superamento di tale concezione è frutto di una rivisitazione del concetto di causalità intesa come congiunzione costante tra cause ed eventi (seguendo la concezione di Hume⁶⁵), fino ad arrivare alla concezione di una causalità come congiunzione probabile⁶⁶ (seguendo la concezione di Suppes⁶⁷).

Quello che rappresenta il *quid pluris* dell'approccio empirico-sperimentale alla realtà non è la capacità di dare risposte ultime a domande ultime. Il valore aggiunto di questa modalità di indagine è di dare risposte accertabili e validabili, rispondendo alla domanda *perché* qualcosa avviene (ricerca della *root cause*) e non solo osservare *che* qualcosa avviene. Per poter dare risposte sul *perché*, è necessario muovere da un modello di indagine che si fonda sull'elaborazione di teorie scientifiche e su un metodo di ragionamento deduttivo-causale.

Se dalle considerazioni generali che precedono, ci spostiamo all'indagine sul reale mediante le tecnologie intelligenti, ci avvediamo che le procedure algoritmiche alla base dei modelli predittivi utilizzano un metodo induttivo. Tale modello si ferma al *sapere che* qualcosa avverrà e si basa sulla correlazione tra eventi e non sulla ricerca causale⁶⁸. Ci muoviamo, quindi, dentro i confini di un ap-

terministico (approccio ontologico), bensì statistico-probabilistico e la conseguente affermazione dell'idea che possono esistere eventi che si verificano "senza una precisa causa", ha decisamente messo in crisi lo stretto determinismo.

⁶⁴ Ricordiamo in questa sede che vari autorevoli autori si sono occupati del tema della causalità. Ad esempio, John Stuart Mill secondo il quale si definisce "causa" la somma di tutte le condizioni positive e negative per il verificarsi dell'evento. Cfr. J.S. MILL, *System of logic ratiocinative and inductive*, London, Longmans, Green & CO, 1886.

Quella concezione ha incontrato problemi quando applicata al mondo delle azioni umane. Infatti, come hanno chiarito Hart e Honoré esiste una famiglia di concetti di causalità di cui quello principale è dato dalla correlazione tra causa ed effetto. Nel caso del comportamento umano, scrivono i due autori, ciò che l'individuo considera causa è un suo intervento, un suo comportamento a cui è riconducibile l'effetto prodotti. Cfr. H.L.A. HART, T. HONORÉ, *Causation in the Law*, Oxford, Clarendon Press, 1985.

⁶⁵ D. HUME, *An Inquiry concerning Human Understanding*, London, A. Millar, 1748.

⁶⁶ C. BARLUCCHI, *Causalità come congiunzione probabile*, in *Quaderni di Sociologia*, 2000, 22, pp. 81-104.

⁶⁷ P. SUPPES, *A Probabilistic Theory of Causality*, Amsterdam, North Holland Publishing Company, 1970.

⁶⁸ Anche l'uomo comune fa largo uso nel suo ragionare sia del metodo deduttivo che del metodo induttivo. Per alcuni autori l'induzione è un'inferenza che porta da premesse particolari a conclusioni generali. Per altri è ogni inferenza che non sia deduzione. Ma l'induzione nel contesto dei ragionatori umani ha mostrato alcuni rilevanti limiti, nonostante sia una forma di ragionamento che, diversamente dalla deduzione, consente di accrescere l'informazione semantica contenuta nelle premesse iniziali. L'induzione «prende le mosse da proposizioni – un insieme di osservazioni e uno sfondo di conoscenze generali e tende a una conclusione che vada oltre il punto di partenza in modo verosimile», cfr. P.N. JOHNSON-LAIRD, *Deduzione Induzione Creatività. Pensiero Umano e Pensiero Macchinico*, cit. p. 95. L'autore menziona anche i rischi del pensiero induttivo umano

proccio che produce risultati che hanno un grado di attendibilità minore rispetto a quelli ottenuti con il metodo deduttivo.

Anderson ha efficacemente definito il passaggio dalla ricerca delle ragioni al sapere che qualcosa accade come *la fine delle teorie*⁶⁹. Non servono ipotesi causali, ma solo correlazioni statistiche.

Anche Paolo Benanti sostiene che «assistiamo a sviluppi tecnologici (capacità di fare) che non corrispondono a nessuno sviluppo scientifico (capacità di conoscere e spiegare)»⁷⁰. Martin Ebers parla di «shift “from causation to correlation”» e sottolinea che «[m]ost data-mining techniques rely on inductive knowledge and correlations identified within a dataset. Instead of searching for causation between the relevant parameters, powerful algorithms are used to spot patterns and statistical correlations»⁷¹. Tale correlazione è intesa come tecnica che interpreta i dati, scovandone regolarità e andamenti. È innegabile che questa tecnica, volta alla ricerca di correlazioni statistiche, produce risultati più velocemente della ricerca di un nesso di causalità. Pertanto, in un mondo votato all'accelerazione dei processi organizzativi e decisionali, il richiamato modello di indagine si sposa perfettamente con le attese economiche e sociali. In tale contesto, è dato per scontato che questo sia il modello da applicare in tutte le attività economicamente e socialmente rilevanti. Detto diversamente, non ci si pone la domanda se esistano alternative e se sia necessario cercarle, soprattutto per settori di attività dove l'affidarsi a modelli predittivi e non causali rischia di avere ricadute negative importanti come, ad esempio, nell'ambito delle scelte pubbliche.

Come è stato di recente osservato «[r]elying solely on predictive models of

partendo da ciò che accadde a Chernobyl. Il disastro fu determinato dal fatto che alcuni «individui si rifiutano di credere a ciò che contraddice le loro più profonde credenze e aspettative» p. 159.

⁶⁹ C. ANDERSON, *The End of Theory*, in *Wired*, 2008, 16, pp. 106-121. Va precisato che quello di cui qui si sta parlando rappresenta un passaggio successivo a quello della *conversione tecnologica della scienza*, utilizzando le parole di Uberto Scarpelli, *Scienza, Sapere, Sapienza*, cit. p. 256. Nell'attuale situazione non si tratta più di tradurre in tecnica la conoscenza scientifica, ma di far procedere la tecnologia senza bisogno di ipotesi teoriche che ne supportino l'uso. Va inoltre sottolineato che la ricerca del solo *sapere che* non induce un significativo ampliamento della conoscenza nella sua funzione più alta. Ecco perché ad esempio «i *Critical Data Studies*, con il supporto della *Computational Social Science*, mettono in discussione il mito di matrice positivista dell'espansione della conoscenza attraverso la costruzione di grandi aggregati di dati e dell'oggettività e della trasparenza delle “informational entities”», cfr. R. PIROSA, *I dilemmi della tecnoregolazione: uno sguardo critico*, in T. CASADEI, S. PIETROPAOLI (a cura di), *Diritto e Tecnologie Informatiche*, Milano, Cedam-Wolters Kluwer, 2021, pp. 275-286, ivi, p. 280.

⁷⁰ P. BENANTI, *The Cyborg: Corpo e Corporeità nell'epoca del post-umano*, Assisi, Cittadella Editrice, 2012, p. 48.

⁷¹ M. EBERS, *Regulating AI and Robotics. Ethical and Legal Challenges*, cit. ivi, p. 45. Si veda anche C.E.A. KARNOW, *The Opinion of Machines*, in W. BARFIELD (eds.), *The Cambridge Handbook of the Law of Algorithms*, Cambridge, Cambridge University Press, 2021, pp. 16-46.

AI in areas as diverse as health care, justice, and agriculture risks devastating consequences when correlations are mistaken for causation»⁷². Ciò accade in quanto è al nesso di causalità che si attribuisce maggiore accuratezza e affidabilità. Pertanto, confondere le correlazioni statistiche con la causalità induce un fraintendimento pericoloso, ossia spinge ad attribuire certezza a proiezioni che in realtà hanno solo un nesso di rilevanza con quanto indagato. Inoltre, il ragionamento induttivo su cui si basano gli algoritmi può essere inficiato da una c.d. fallacia di generalizzazione indebita. Questo accade quando il campione di dati su cui è addestrata l'IA non è sufficiente per operare una proiezione, ma questa viene lo stesso effettuata.

Per far fronte ai problemi e ai limiti⁷³ che presenta l'indagine del reale, prevalentemente, se non esclusivamente fondata su modelli induttivi di correlazioni statistiche, ci si sta muovendo per integrare l'indagine causale⁷⁴ negli studi sull'IA. La ricerca nel settore della *causal AI* sembra avere già dimostrato il suo potenziale in alcuni settori⁷⁵.

Sebbene questo campo di ricerca presenti problemi pratici, in particolare quelli che riguardano la capacità dell'IA di cogliere i nessi causali, simulando il ragionamento causa-effetto tipico dell'essere umano, si tratta di problemi tecnici che, nel tempo e con nuove conoscenze, potranno essere risolti. Già molto si sa circa le questioni che vanno tenute in considerazione per evitare problemi e rischi di un'IA causale: in relazione all'uso, alla raccolta e alla qualità dei dati; in relazione alla disponibilità di *open-source tools* per analisi del codice sorgente e restituzione di feedback; in relazione al miglioramento necessario per valutare la prestazione algoritmica ecc. Si tratta, quindi, di una frontiera già in esplorazione di cui appare opportuno evidenziare alcune ricadute simboliche e assiologiche importanti. Innanzitutto, muovere verso una IA causale consente, se usata *per il bene*⁷⁶, di rimettere al centro il *perché qualcosa accade* o accadrà e non solo *che*

⁷²S.K. SGAIER, V. HUANG, G. CHARLES, *The Case for Causal AI*, in *Stanford Social Innovation Review*, 2020, 18, 3, pp. 50-55, accessibile al link https://ssir.org/articles/entry/the_case_for_causal_ai, ultimo accesso dicembre 2021.

⁷³Z. OBERMEYER, B. POWERS, C. VOGELI, S. MULLAINATHAN, *Dissecting racial bias in an algorithm used to manage the health of populations*, in *Science*, 2019, 366, 6464, pp. 1-7. L'articolo tratta dei bias razziali a seguito dell'impiego di un algoritmo nell'ambito della salute.

⁷⁴B. SCHÖLKOPF, F. LOCATELLO, S. BAUER, N.R. KE, N. KALCHBRENNER, A. GOYAL, Y. BENGIO, *Towards Causal Representation Learning*, in *Special Issue of Proceedings of the IEEE-Advances in Machine Learning and Deep Neural Networks*, 2021, pp. 1-23.

⁷⁵Nel campo del cambiamento climatico è stata utilizzata per comprendere se e come gli esseri umani contribuiscono alle cause, per altri campi di applicazione cfr. S.K. SGAIER, V. HUANG, G. CHARLES, *The Case for Causal AI*, cit.

⁷⁶Vogliamo chiarire che la nozione non indica un bene oggettivo, predefinito, coerente con un ordine dato. L'espressione indica molto più umilmente la scelta di orientare gli usi della tecnologia per il miglioramento delle condizioni di vita di tutti, esseri umani e pianeta terra, in una di-

qualcosa accadrà. Se la domanda sulle cause torna a guidare la conoscenza, questo avrà una ricaduta positiva anche su quelle strategie politiche volte a mantenere l'essere umano al centro del progresso tecnologico. È innegabile che l'approccio incentrato sulla ricerca delle cause ha riflessi importanti sul pensiero e sull'autocomprensione di sé stessi. È questa la modalità di indagine che consente di evidenziare l'esistenza di una *costruzione intenzionale dell'esperienza*⁷⁷, allontanandosi così dall'illusione che esista un «ordine di cose [...] senza la partecipazione della nostra riflessione e dei nostri intenti», evitando così di «trasformare la realtà in mito e privare della propria ragion d'essere tutta la sfera vitale dell'attività impegnata e riflessiva»⁷⁸.

È di fondamentale importanza non confondere l'approccio scientifico moderno, nei sensi richiamati, con altre forme di conoscenza attuali come, ad esempio, il *Machine Learning (ML)*. Come si è osservato, il ML «is a form of numerical pattern finding with predictive power, prompting comparisons with science. But rather than being universal and objective, it produces knowledge that is irrevocably entangled with specific computational mechanisms and the data used for the training», quindi «[t]here is no larger explanatory structure to fall back on because the outcomes are based on large-scale correlations, not on logical causation. Thus, there is no cumulative support for the proposed explanation from other non-falsified hypotheses, as there is in the physical sciences»⁷⁹. In altre parole, in certi campi come quello della salute e della giustizia, è utile e opportuno poter contare sul procedimento di falsificabilità, per evitare di trarre false conclusioni. Va, quindi, proseguito lo studio per un'IA causale. Muovere verso un'IA causale significa un *ritorno alle teorie* e un passo verso il superamento dell'oggettivizzazione degli individui, che è una deriva grave dell'uso dell'IA, in particolare quando questa serve a profilare i soggetti.

Nello studio del Consiglio d'Europa sul tema della responsabilità e dell'intelligenza artificiale si chiarisce, infatti, che l'uso di tecnologie per personalizzare prodotti e servizi tende a oggettivizzare gli individui, che, a causa di ciò, non sono più visti come agenti morali. Questa oggettivizzazione si realizza in quanto «individuals are singled out, not on the basis of any causal theory, but simply on the basis of correlations in data sets. As a result, these systems typically do not provide any reasoned account to individuals explaining why they have been singled out for treatment of a particular kind. Secondly, their underlying logic and their processing operations are highly complex and opaque, in ways that are

mensione di eguaglianza e rispetto della dignità, secondo le definizioni che di questi principi se ne dà nell'attuale momento storico nel contesto europeo.

⁷⁷J. DEWEY, *La teoria sperimentale della conoscenza*, in R. FREGA (a cura di), *Logica Sperimentale, Teoria Naturalistica della Conoscenza e del Pensiero*, Macerata, Quodlibet, 2008, pp. 77-95, ivi, p. 94.

⁷⁸*Ibidem*, p. 95.

⁷⁹D. MCQUILLAN, *People's Councils for Ethical Machine Learning*, in *Social Media+Society*, 2018, 4, 2, pp. 1-18.

practically, and sometimes technically, incomprehensible [...]. In other words, because many contemporary ML systems are designed to capture, commodify and optimise value extraction in the interests of the system owner, by tracking and analysing from the digital traces of the daily behaviour of individuals, they are not primarily concerned with identifying the *reasons* why individuals behave in particular ways»⁸⁰. Muovendo da queste osservazioni, è possibile ricollegarsi a quanto si diceva qualche pagina addietro, ossia che la diffusione del metodo di indagine fondato sull'induzione fa il gioco di chi vuole costruire un preciso modo di guardare al mondo, ovvero il più possibile slegato dalla ricerca di risposte sul perché. Di fatto è l'indagine sulle cause degli accadimenti e dei comportamenti che apre l'opzione di scelta, perché, spiegando le ragioni di un dato evento, consente anche di decidere se, avendone la possibilità, sia utile intervenire per modificarlo in futuro o meno. Questa apertura verso il cambiamento è scomoda per chi vuole, invece, che gli individui seguano modelli selezionati di comportamento. In particolare, in relazione a pratiche ben precise, di carattere economico-commerciale, si è infatti parlato di «*interested readings of reality*»⁸¹. Da ciò consegue che l'utilità di impiegare i sistemi di IA per *sapere che*, senza interessarsi di *sapere perché*, è prevalentemente per certi gruppi di attori sociali ed economici. A questa osservazione va aggiunto che la rappresentazione del reale effettuata su questi presupposti è parziale, spesso faziosa e più manipolabile di quella fondata su un metodo di indagine deduttivo-causale. In sintesi, per mantenere un maggiore controllo sulla lettura della realtà appare indispensabile sviluppare una tecnologia intelligente in grado di fornire ragioni dei risultati che produce.

Se risulta, pertanto, auspicabile continuare gli studi per lo sviluppo di un'IA causale, è già tuttavia possibile anticipare alcuni rischi anche di questo modello. Quando si usa il principio di causalità per spiegare i comportamenti umani, vi è il rischio di una lettura deterministica dell'agire umano. Se l'IA dovesse fornire dati, ad esempio nell'ambito delle neuroscienze, utili a supportare tale visione, si potrebbe essere indotti a ridurre la spiegazione dei comportamenti umani alle sole cause fisiche, genetiche e neurologiche, rimettendo di nuovo in discussione le categorie della libertà e della responsabilità.

Per evitare rischi e criticità di un'IA causale occorre porre delle regole al suo impiego, per evitare che le tecnologie intelligenti divengano tecnologie «del potere, che regolano la condotta degli individui e li assoggettano a determinati scopi o domini esterni, dando luogo a una oggettivazione del soggetto»⁸².

⁸⁰ COUNCIL OF EUROPE STUDY DGI (2019)05, *Responsibility and AI*, p. 36, <https://rm.coe.int/responsability-and-ai-en/168097d9c5>.

⁸¹ B. RIEDER, *Big Data and the Paradox of Diversity*, in *Digital Culture and Society*, 2016, 2, 2, pp. 39-54, *ivi*, p. 41.

⁸² M. FOUCAULT, *Tecnologie del sé*, (a cura di) L. MARTIN, H. GUTMAN, P.H. HUTTON, Torino, Bollati-Boringhieri, 1992, p. 13.

Nonostante i richiamati potenziali pericoli, vale tuttavia la pena di percorrere la strada di una ricerca nel campo dell'IA causale, perché è quella più vicina a quella forma di sapere che, andando a fondo delle ragioni di ciò che accade, consente di approfondire la conoscenza. Si tratta di una conoscenza che, spiegando *perché qualcosa accade*, non fermandosi a descrivere *che qualcosa accade*, apre gli orizzonti di scelta individuali e collettivi. In tal modo, si pongono le basi per cambiare e modificare i comportamenti in funzione della realizzazione di programmi sociali, volti a tutelare la giustizia sociale e la dignità individuale.

Ciò appare particolarmente importante, se non urgente, in un contesto storico in cui si è venuta evidenziando la necessità di superare egoismi dettati da interessi particolari e modalità di azione limitate al presente per, invece, agire in modo che le conseguenze della nostra «azione siano compatibili con la permanenza di un'autentica vita umana sulla terra»⁸³.

5. La questione tecnica e la semantica dell'antropomorfizzazione: dominio e potere sull'individuo

Per riassumere quanto osservato nel paragrafo precedente, possiamo prendere in prestito le parole di Frank Pasquale «chi si occupa di etica e di tecnologia si rende sempre più conto che i principali problemi dell'AI non riguardano gli aspetti tecnici ma il suo ruolo sociale»⁸⁴. Tale ruolo si realizza lungo direttrici diverse, ma interconnesse. Per individuare tali direttrici possiamo mutuare i quattro tipi fondamentali di tecnologia elaborati da Michel Foucault⁸⁵. Innanzitutto, l'impatto dell'IA riguarda le tecnologie della produzione volte a realizzare e trasformare oggetti. In secondo luogo, l'IA si è impadronita delle tecnologie che consentono di far uso del linguaggio (segni, significati, simboli ecc.)⁸⁶. In terzo luogo, le tecnologie intelligenti sono tecnologie del potere. Infine, l'IA è usata nelle c.d. tecnologie del sé, che secondo Foucault, permettono di «eseguire, coi propri mezzi o con l'aiuto degli altri, un certo numero di operazioni sul proprio corpo e sulla propria anima»⁸⁷.

In tale senso, si muove il dibattito sul biopotenziamento umano di cui par-

⁸³ H. JONAS, *Il principio di responsabilità*, cit. p. 16.

⁸⁴ F. PASQUALE, *Le nuove leggi della robotica*, cit. p. 158.

⁸⁵ M. FOUCAULT, *Tecnologie del sé*, cit.

⁸⁶ Pensiamo alle reti neurali come LaMda che sono in grado di sostenere conversazioni complesse con gli esseri umani, tanto che di recente vi è chi ha considerato questa IA senziente. La vicenda è ben narrata nell'articolo di Khari JOHNSON dal titolo *LaMda e la trappola dell'intelligenza artificiale senziente* del 18 giugno 2022 pubblicato su *Wired.it* e accessibile a questo link <https://www.wired.it/article/intelligenza-artificiale-lamda-dibattito-ai-senziente/>.

⁸⁷ M. FOUCAULT, *Tecnologie del sé*, cit. p. 13.

remo nel capitolo quarto. Qui ci limitiamo ad accennare che tale dibattito riguarda l'uso di mezzi farmacologici e non farmacologici, come le neurotecnologie (es. Brain-Computer Interfaces, BCIs), per migliorare o potenziare capacità cognitive e fisiche di chi è sano. Le ricadute sociali delle neurotecnologie, che si appoggiano all'IA, sono sia positive che negative. Le ricadute positive riguardano il miglioramento della qualità della vita attraverso una migliore comprensione delle potenzialità del nostro cervello. Le ricadute negative si riferiscono all'abuso di queste neurotecnologie per controllare e manipolare, in maniera eterodiretta, il nostro cervello. Saremo confrontati con un paradosso di questo tipo: il nostro cervello è controllato dall'IA che è però insipiente, in grado di *agere sine intelligere*⁸⁸, ma che, in mano a chi fa del controllo sugli altri il fine delle proprie azioni, diviene un formidabile strumento di ridefinizione dell'ordine sociale. Come ha scritto Laurent Alexandre «[...] i sorprendenti progressi compiuti di recente nella lettura del cervello possono farci temere la dissoluzione della nostra individualità in un enorme "hub" globale delle coscienze [...] C'è il rischio che i poteri politici si rivelino particolarmente propensi all'estensione delle tecniche di trasmissione del pensiero [...] il prezzo da pagare per questa sicurezza, ovviamente, sarebbe il controllo politico assoluto. La democrazia perfetta, in cui ogni cittadino è in costante comunicazione con la mente dei suoi dirigenti, sarebbe anche la morte della democrazia»⁸⁹.

Di fatto, sotteso al controllo sul cervello vi è il potere sulla narrazione dell'autonomia, della libertà, della (im)prevedibilità, vale a dire su quelle nozioni che formano la tassonomia concettuale della narrazione di un'IA antropomorfa per eccesso, di cui si è dato conto nel primo capitolo.

6. Il problema tecnico che dissimula un problema socioculturale

Da più voci⁹⁰ si leva il grido di pericolo del penetrare nelle menti delle persone della convinzione che si possano giustificare scelte politiche sul piano so-

⁸⁸ È un'espressione utilizzata da Luciano Floridi.

⁸⁹ L. ALEXANDRE, *La guerra delle intelligenze. Intelligenza artificiale contro intelligenza umana*, cit. p. 255.

⁹⁰ Cfr. E. SADIN, *Critica della ragione artificiale. Una difesa dell'umanità*, Roma, LUISS University Press, 2019; S. ZUBOFF, *Il capitalismo della sorveglianza*, cit.; M. CHIRIATTI, *Incoscienza artificiale*, cit.; F. PASQUALE, *Le nuove leggi della robotica*, cit.; T. KARPPI, "The Computer Said So": *On the Ethics, Effectiveness, and Cultural Techniques of Predictive Policing*, in *Social Media+Society*, 2018, April-June, 4, 2, pp. 1-9. Al tema in questione, il *New York Times* ha dedicato già nel 2015 (19 novembre) un dibattito intitolato *Can predictive policing be ethical and effective?*, cfr. in quel dibattito in particolare la posizione di S.P. GANGADHARAN, *Predictive Algorithms are not Inherently Unbiased*, accessibile al link <https://www.nytimes.com/roomfordebate/2015/11/18/can-predictive-policing-be-ethical-and-effective/predictive-algorithms-are-not-inherently-unbiased>. Cfr. anche J.C. MAGLHÃES, *Do Algorithms Shape Character? Considering Algorithmic Ethical Subjectivation*, in *Social Media+Society*, 2018, April-June, 4, 2, pp. 1-10.

ciale, criminale, economico e occupazionale semplicemente perché “lo ha detto il computer o l’algoritmo”. Insomma, si è forse già insinuato il dubbio che non siano più gli umani a prendere decisioni, ma lo facciano le macchine e i sistemi intelligenti. In breve, si sta dimenticando che «il mondo, oltre a essere plasmato dall’evoluzione naturale, è creato attivamente dallo sviluppo culturale di noi umani, con il supporto della tecnologia»⁹¹. Ciò accade per diversi motivi, ma l’accento costantemente posto sugli aspetti tecnici, o meglio sul mandato esclusivamente tecnico delle trasformazioni in atto, gioca un ruolo centrale nell’offuscare l’esistenza della componente umana. Per spiegare questo aspetto si può ricorrere alla concezione espressa attraverso il termine *Kulturtechniken*, che ha trovato la traduzione inglese in *cultural techniques*⁹². Con tale nozione si indica «operative chains composed of actors and technological objects that produce cultural orders and constructs which are subsequently installed as the basis of these operations. At the core of [...] cultural techniques is the notion that fairly simple operations coalesce into complex entities which are then viewed as the agents or sources running these operations»⁹³. E in un qualche modo, non sempre chiaramente spiegabile, le catene procedurali e le tecniche di collegamento producono nozioni ed oggetti che sono caricate di *essentialized identities*⁹⁴. Seguendo questa linea di pensiero si può comprendere come si formano quelle convinzioni che portano a credere che oggi giorno siano gli algoritmi ad assurgere a decisori autonomi, siano essi impiegati nell’ambito della giustizia, della medicina o di altre attività in cui trovano spazio⁹⁵. Ragionando nei termini dati dalla definizione di *cultural techniques*, è possibile compiere due operazioni. La prima operazione è quella di ragionare in termini di tecniche culturali e non di tecnologia, mettendo così in risalto l’aspetto di costruzione della realtà con una forte componente soggettiva, che consente di respingere i dogmi della neutralità e dell’oggettività. La seconda operazione è quella di evidenziare che questa costruzione avviene avvalendosi di distinzioni realizzate sulla base dell’analisi di dati, campioni statistici, mai completi né totali, né tantomeno rappresentativi di

⁹¹ M. CHIRIATTI, *Incoscienza artificiale*, cit. p. 15.

⁹² G. WINTHROP-YOUNG, *Cultural techniques: Preliminary remarks*, in *Theory Culture and Society*, 2013, 30, 6, pp. 3-19.

⁹³ G. WINTHROP-YOUNG, *Discourse, media, cultural techniques: The complexity of Kittler*, in *MLN*, 2015, 130, 3, pp. 447-465, ivi, p. 458.

⁹⁴ *Ibidem*.

⁹⁵ È certamente condivisibile quanto scrive Fioriglio sul tema, ossia che «[...] il riferimento, oramai di prassi, all’algoritmo abbia certamente il pregio di evidenziare in modo fortemente evocativo e dirompente il passaggio epocale dalla mera automazione nell’acquisizione e nel trattamento delle informazioni a una automazione “intelligente”, grazie all’evoluzione dell’intelligenza artificiale [...]», cfr. G. FIORIGLIO, *La società algoritmica fra opacità e spiegabilità: profili informatico-giuridici*, in *Ars Interpretandi*, 2021, 1, pp. 53-67, ivi, p. 54. Cfr. anche V. BARASSI, *I figli dell’algoritmo. Sorvegliati, tracciati, profilati dalla nascita*, Roma, LUISS University Press, 2021.

tutto il reale⁹⁶. Questa analisi si avvale dei significati con cui il sistema di IA è stato alimentato. Si tratta di significati determinati storicamente, in quanto le predizioni dell'IA si basano necessariamente su dati passati sui quali costruiamo la predizione futura, dati necessitanti di delimitazione e determinazione semantica. Ciò perché la costruzione intenzionale del reale e dell'esperienza, «in ogni tipo di anticipazione, sia essa artistica, utilitaria o tecnologica, sociale o morale»⁹⁷ rischia altrimenti di non avere il successo auspicato e di non raggiungere lo scopo previsto. In questa cornice, risulta comprensibile l'esigenza di superare l'opacità algoritmica, nei limiti del possibile, e di garantire trasparenza. Non si tratta di garantire una trasparenza totale a tutti i livelli di comunicazione ed utilizzo. Ma piuttosto di fornire quelle informazioni sulla trasparenza che sono utili nella situazione e nel caso concreto, in modo che chi utilizza un sistema, a seconda che si tratti di un produttore o di un consumatore finale, abbia a disposizione gli strumenti che sono utili, nel suo caso, a far valere forme di responsabilità, quando vi sono problemi e danni. In questa direzione va la proposta di Margot Kaminski di creare una tassonomia della trasparenza per un'*accountable* IA⁹⁸. Diversamente da molti autori che propongono di creare forme e gradi di trasparenza, l'autrice distingue tra due forme principali di trasparenza, ossia una trasparenza individualizzata e una sistemica. La prima, *individualized transparency*, riguarda sia la predisposizione di modalità di informazione all'individuo con lo scopo di proteggerlo e accrescerne l'autonomia sia la possibilità di rendere visibile e ricostruibile la logica alla base del sistema per poterlo giustificare⁹⁹. La seconda, *systemic transparency*, mira a evidenziare errori, *bias* e discriminazioni sia nel sistema macchina che in quello umano che lo dovrebbe controllare, per poter intervenire in modo mitigante e correttivo. È una trasparenza indirizzata a diversi attori, ma che difficilmente interessa il grande pubblico¹⁰⁰. L'idea di modulare la trasparenza in base al destinatario finale dell'informazione può avere sicuramente il vantaggio di non appesantire eccessivamente l'utente finale con informazioni, per lui superflue e che magari fatica a gestire e comprendere, allo stesso tempo però non deve diventare un'euristica per aggirare gli obblighi informativi e mantenere vivo l'atteggiamento tecnocentrico. Il tema della trasparenza richiede che si tengano presenti molti aspetti. La trasparenza algoritmica dipende da una combinazione di processi e da proprietà tecniche del sistema e, per tale motivo, presenta sfide importanti. In ogni caso, la trasparenza ha come

⁹⁶ Cfr. L. GITELMAN, *Raw Data is an Oxymoron*, Cambridge, MIT Press, 2012.

⁹⁷ J. DEWEY, *La teoria sperimentale della conoscenza*, cit. p. 94.

⁹⁸ M. KAMINSKI, *Understanding Transparency in Algorithmic Accountability*, in W. BARFIELD (eds.), *The Cambridge Handbook of the Law of Algorithms*, Cambridge, Cambridge University Press, pp. 119-138, ivi, p. 127.

⁹⁹ *Ibidem*, p. 129.

¹⁰⁰ *Ibidem*.

primario obiettivo quello di rendere possibile la responsabilità, nel senso di *accountability*. In questa direzione si era già espresso il *Panel for the future of science and technology (STOA)* nel 2019¹⁰¹. Il *Panel* sottolinea come la trasparenza sia una questione primariamente tecnica, mentre quella dell'*accountability* una questione etico-giuridica. Occorre, quindi, a questo punto muoversi più addentro alle questioni giuridiche per vedere come le questioni tecniche possano essere affrontate dal diritto.

¹⁰¹ PANEL FOR THE FUTURE OF SCIENCE AND TECHNOLOGY (STOA), *A governance framework for algorithmic accountability and transparency*, 2019, accessibile al link [https://www.europarl.europa.eu/stoa/en/document/EPRS_STU\(2019\)624262](https://www.europarl.europa.eu/stoa/en/document/EPRS_STU(2019)624262). Per un'analisi del documento cfr. M. SAPORITI, *Questioni di "intelligenza politica". Prospettive europee in materia di Intelligenza Artificiale e di proceduralità algoritmica*, in *Notizie di Politeta*, 2021, 143, pp. 87-103.

CAPITOLO III

PROGRESSO SCIENTIFICO E TECNOLOGICO E QUESTIONE GIURIDICA: LA SEMANTICA DEL *LICERE*

SOMMARIO: 1. Introduzione. – 2. Che rapporto è possibile tra diritto, scienza e tecnologia? – 2.1. Tecnologia come oggetto di regolazione. – 2.2. Tecnologia come agente regolatore. – 3. Categorie giuridiche alla prova dell'IA. – 4. I diritti fondamentali sono la cornice adeguata a regolare l'innovazione tecnologica? Il caso del riconoscimento facciale. – 4.1. Il principio di precauzione e la tutela effettiva dei diritti. – 4.2. Il principio di esplicabilità e i diritti portati a compimento.

1. *Introduzione*

Nei capitoli che precedono, abbiamo fatto il punto sulla complessità della narrazione semantica dell'IA, che si muove tra una radicata dialettica degli opposti e l'ideologia antropomorfizzante. Abbiamo altresì richiamato le problematiche che derivano da concezioni deterministiche dello sviluppo tecnologico. Nell'affrontare gli aspetti linguistici, che agevolano la strumentalizzazione della narrazione sulle tecnologie contemporanee definite intelligenti, abbiamo focalizzato l'attenzione sulle trappole linguistiche. Inoltre, sul piano tecnico abbiamo sollevato il velo che copre posizioni etiche ben precise sulla spettanza delle decisioni relative al percorso dell'innovazione tecnologica, posizioni che si celano dietro la concezione di un esclusivo mandato tecnico-scientifico.

In questo capitolo, in ideale prosecuzione con i capitoli precedenti, ci sposteremo a considerare la declinazione giuridica della narrazione dell'IA. L'intento è quello di far emergere le implicazioni della dialettica degli opposti, dell'ideologia antropomorfizzante, e della retorica oggettivistica sulla dimensione giuridica di questo fenomeno. Per procedere in tale direzione, è necessario, innanzitutto, richiamare preliminarmente il controverso rapporto tra diritto e tecnologia. Ci soffermeremo poi su aspetti più strettamente pertinenti la regolazione giuridica dell'IA. Ai fini di una specifica definizione dei profili di problematicità giuridica della narrazione ideologica dell'IA, prenderemo in considerazione le categorie maggiormente esposte all'influenza della dialettica degli opposti, per via dell'uso ubiquo – nel linguaggio ordinario e in quello giuridico – delle nozioni che espri-

mono tali categorie. Si pensi, ad esempio, alla categoria della personalità giuridica e al correlato concetto di persona.

Infine, rifletteremo sull'adeguatezza dei diritti fondamentali come misura e limite di ciò che l'innovazione tecnologica consente di fare.

2. Che rapporto è possibile tra diritto, scienza e tecnologia?

Nell'attuale vivace discussione a livello istituzionale europeo, uno degli argomenti affrontati è quello del rapporto tra tecnologia e diritto.

La domanda a cui si tenta di trovare una risposta è formulata come segue: *What if law shaped technology? What if technologies shaped the law?* In questi termini, la questione è analizzata nella rassegna di articoli pubblicati nella serie mensile *What if* del *Panel for the Future of Science and Technology (STOA)* del Parlamento europeo, espressamente dedicata a queste tematiche¹. Come si può osservare, i titoli pongono la questione in forma di alternativa, semplificando di molto il problema delle complesse interazioni tra diritto e tecnologia. In realtà, quello su cui ci si interroga è in che misura la tecnologia determina la direzione in cui si deve muovere il diritto e, d'altro canto, in che misura la tecnologia può essere oggetto di regolazione giuridica. Questi interrogativi si sono posti anche nella definizione del rapporto tra scienza e diritto².

Come noto, la storia del diritto è in buona parte costruita attorno alla ricerca dei caratteri differenziali tra ciò che è giuridico e ciò che non lo è. A questa ricerca ha dato un centrale contributo la filosofia del diritto di indirizzo analitico linguistico³. La filosofia del diritto, praticata in tale prospettiva, ha contribuito ad analizzare e chiarire, per un verso, i caratteri differenziali del diritto da altre realtà normative, e per altro verso, a porre le basi per affrontare razionalmente la c.d. *is-ought question*.

Detto diversamente, il chiarimento operato dalla filosofia del diritto di indirizzo analitico linguistico ha contribuito, all'interno della sfera del normativo, a individuare i *limiti epistemici*⁴, ossia quelli che «differenziano i contorni del diritto dagli altri sistemi normativi con cui convive all'interno dello spazio socia-

¹ *What if law shaped technology?; What if technologies shaped the law?; What if technologies had their own ethical standards?; What if technologies challenged our ethical norms?* Tutti reperibili al seguente link: <https://www.europarl.europa.eu/stoa/en/indexsearch?ordering=relevance&query=at+a+glance&scope=CURRENT>.

² Cfr. P. BORSELLINO, *La prospettiva della filosofia del diritto*, in B. LIBERALI e L. DEL CORONA (a cura di), *Diritto e Valutazioni Scientifiche*, Torino, Giappichelli, 2022, pp. 37-53.

³ Nel contesto italiano, tale orientamento ha trovato originale impulso negli studi di Norberto Bobbio e Uberto Scarpelli.

⁴ F.J. ANSUÁTEGUI ROIG, *Norme, Giudici, Stato Costituzionale. Frammenti di un Modello Giuridico*, Torino, Giappichelli, 2020, p. 91.

le»⁵, rispetto ad altri contesti normativi presenti nell'ambiente sociale⁶. La questione dei limiti del giuridico è stata al centro della battaglia giuspositivistica per la separazione tra diritto e morale, volta a garantire autonomia al diritto positivo rispetto al c.d. diritto naturale⁷. Ma la filosofia del diritto nell'accezione richiamata ha altresì proficuamente aiutato a delimitare il rapporto tra *diritto e natura*⁸ – tra *dover essere* ed *essere* –, mettendo in luce due livelli discorsivi: uno in funzione assertivo-descrittivo e l'altro in funzione prescrittivo-valutativo⁹. Questa distinzione tra le funzioni del linguaggio rappresenta il presupposto per meglio comprendere il rapporto tra scienza e diritto. La prima, infatti, si esprime mediante asserzioni, passibili di controllo di verità e/o falsità. Il secondo si esprime mediante prescrizioni, che non sono vere o false, ma hanno lo scopo di orientare i comportamenti nella direzione auspicata. La distinzione sul piano linguistico consente di far emergere la diversità dei metodi e degli scopi della scienza, da una parte, e del diritto, dall'altra. La scienza costruisce, tramite il metodo sperimentale, ipotesi su eventi, fenomeni e comportamenti, che verranno corroborate attraverso il controllo intersoggettivo e spiega i fenomeni con richiamo al principio di causalità. Il diritto regola le condotte umane alla luce del principio di imputazione¹⁰. Le regole giuridiche sono la specificazione dei principi e dei valori ad essi sottesi e assunti come rilevanti in un dato momento storico, con lo scopo di garantire un certo ordine sociale e una convivenza pacifica tra i consociati. A fronte delle cruciali differenze metodologiche, così come degli obiettivi e degli scopi perseguiti tra diritto e scienza, si è posto il problema se e in che misura la scienza possa essere oggetto di regolazione giuridica e, al contempo, quale contributo essa possa fornire al diritto.

Per rispondere al primo quesito, occorre fare una precisazione fondamentale. Chiedersi se la scienza possa essere oggetto di regolazione giuridica non significa sollevare dubbi sulle procedure e sui metodi sperimentali alla base del progresso scientifico. La questione, quindi, non riguarda mai una valutazione sul metodo sperimentale con cui la scienza si è evoluta. Non è questo un compito

⁵ *Ibidem*.

⁶ M. JORI, A. PINTORE, *Manuale di teoria generale del diritto*, Torino, Giappichelli, 1995, p. 36. La ricerca ha pertanto riguardato i caratteri differenziali del diritto «rispetto a concetti e fenomeni affini, e con esso confondibili, come la morale e i costumi sociali».

⁷ Per alcuni riferimenti di base, cfr. H. KELSEN, *Lineamenti di dottrina pura del diritto*, Torino, Einaudi, 1952.

⁸ Alcuni riferimenti essenziali: H. KELSEN, *Lineamenti di dottrina pura del diritto*, cit.; H.L.A. HART, *Il concetto di diritto*, Torino, Einaudi, 1965; H.L.A. HART, *Il positivismo e la separazione tra diritto e morale*, in ID., *Contributi all'analisi del diritto*, Milano, Giuffrè, 1964, pp. 105-166; N. BOBBIO, *Giusnaturalismo e positivismo giuridico*, Roma-Bari, Laterza, 2011. Cfr. anche W. OTT, *Der Rechtspositivismus*, Berlino, Duncker & Humblot, 1992.

⁹ Abbiamo dettagliato queste funzioni nel precedente capitolo.

¹⁰ H. KELSEN, *Lineamenti di dottrina pura del diritto*, cit.

che spetta al diritto. Quello che può essere oggetto di valutazione etico-giuridica sono la destinazione e gli usi dei risultati ottenuti con quel metodo. Tale valutazione è non solo possibile, ma anche dovuta, se, rifiutando logiche oggettivistiche e deterministiche, si tiene presente che il sapere scientifico si colloca nello spazio sociale e culturale in cui vivono i suoi promotori. In poche parole, le scelte circa quali ricerche incentivare e quali risultati impiegare tra quelli ottenuti includono valutazioni etiche, sociali, culturali e giuridiche, oltre a quelle strettamente scientifiche. La valutazione etica sulla sostenibilità dell'innovazione scientifica e la conseguente regolazione giuridica si rendono necessarie perché ogni passo, che la scienza compie, apre orizzonti di scelta inediti¹¹. Questa eccedenza, che deriva da nuove conoscenze sul funzionamento di meccanismi naturali, rappresenta lo scarto tra ciò che si è fatto fino a un certo momento e ciò che è opportuno fare da lì in poi. In poche parole, la scienza ci dice *che qualcosa è in un certo modo*, mentre è sul piano etico-normativo che decidiamo che quella conoscenza va indirizzata verso il perseguimento di certi scopi, abbandonandone altri. L'influenza delle conoscenze scientifiche sui sistemi normativi è indiretta, ma profonda, in quanto rimette costantemente in gioco le domande di fondo sui valori a cui ispirare l'uso dei risultati e delle conversioni tecniche per migliorare l'esistenza individuale e collettiva.

Queste valutazioni sono certamente ispirate dalla scienza, in quanto i suoi risultati consentono di eliminare certe precondizioni influenzate da pregiudizi, come, ad esempio, superare fraintendimenti sul funzionamento di certi processi biologici e sgombrare il campo da inutili e fuorvianti luoghi comuni. Ma sono valutazioni non direttamente derivate dal dato scientifico, perché per poter essere utili sul piano delle politiche sociali, culturali e criminali richiedono un confronto ulteriore sui valori che ispirano le scelte in un dato momento storico, entro un preciso progetto politico-giuridico. Chi sostiene questa linea di pensiero si batte, per un verso, per il riconoscimento della possibilità di un discorso etico-giuridico razionale¹², e per altro verso, contro l'appiattimento del valore sul

¹¹ Si pensi alla possibilità di correggere il genoma umano con la tecnica CRISP-CAS9 e alle sue ricadute sul piano sociale e individuale come ben ricordato da S. JASANOFF, G.B. HURLBUT, K. SAHA, *CRISPR Democracy. Gene Editing and the Need for Inclusive Deliberation*, in *Issues in Science and Technology*, 2015, 32, 1, pp. 25-32.

¹² Quello di riconoscere significanza agli usi linguistici normativi, ad esempio, ai giudizi di valore è stato un problema giusfilosofico molto importante. Se i giudizi di valore sono significanti, ciò consente di ammetterli all'argomentazione razionale e accettare che possono svolgere un rilevante ruolo di guida dei comportamenti, influenzando le scelte in ambito pubblico. Sul tema cfr. M. De CARO, D. MACARTHUR (a cura di), *Hilary Putnam. La filosofia nell'età della scienza*, Bologna, Il Mulino, 2012, p. 186, Hilary Putnam esemplifica bene questo aspetto con riguardo alle opinioni espresse dall'economista Lionel Robbins, che, durante la fase culminante della Depressione, «argomentò contro l'idea complessiva di redistribuzione dei redditi per la ragione filosofica che i giudizi di valore (secondo lui) sono completamente al di fuori della sfera della ragione».

fatto, a cui è sottesa l'idea di un'etica non ontologicamente fondata, bensì costruita sulle scelte responsabili dei singoli e della collettività.

Quanto sinora osservato circa il rapporto tra scienza e diritto risulta propeudeutico ad affrontare la relazione tra tecnologia e diritto, che interessa il nostro percorso di analisi. Nel dibattito istituzionale europeo, la relazione tra tecnologia e diritto è spesso posta in termini di alternativa tra la tecnologia come oggetto di regolazione e la tecnologia come agente regolatore¹³. Affronteremo di seguito la prima questione, dedicando alla seconda il successivo paragrafo.

2.1. Tecnologia come oggetto di regolazione

Trattare il tema della tecnologia come oggetto di regolazione dal punto di vista istituzionale europeo significa concentrarsi su due questioni così riassumibili: se e in che modo il diritto possa regolare i prodotti dell'innovazione tecnologica. Vengono analizzate due tipologie di intervento. La prima è definita reattiva o *ex-post facto*, mentre la seconda possiamo definirla *ex-ante facto*, preventiva o anticipatoria. L'intervento "reattivo" rappresenta la modalità tradizionale con cui il diritto interviene. Si cerca nelle norme esistenti una risposta per orientare i comportamenti che hanno già espresso i loro effetti, ma che possono configurarsi come fattispecie nuove, non direttamente sussumibili sotto regole preesistenti.

Di norma è l'opera interpretativa dei giudici che consente di definire fattispecie nuove, non ancora espressamente previste dalle regole esistenti.

Nell'era tecnologica la modalità di intervento reattivo è, spesso, criticata e considerata insufficiente. Ciò perché si dice: l'intervento giuridico si realizza a danno già causato e non è utile ad arginare l'accelerazione esponenziale dello sviluppo tecnologico.

A fronte di questi limiti, ci si è mossi per individuare strumenti giuridici adatti a stare al passo con la velocità con cui la tecnologia si sviluppa.

Si è puntato a rimedi in grado di prevenire potenziali danni o complessi conflitti di interesse. Tra questi rimedi si annoverano le *sunset clauses*, che consentono di rivedere la normativa dopo alcuni anni dalla sua entrata in vigore per valutare la necessità di modificarla; o l'*horizon scanning*, che consente di anticipare rischi e opportunità; o l'*innovation deal*, pensato per sostenere le imprese innovative, evitando che la loro attività si blocchi di fronte alle richieste di adeguamento alle normative esistenti; e le *regulatory sandboxes*, che consentono alle aziende di accedere a un ambiente di sperimentazione controllato, a contatto con il mercato reale, ma con meno vincoli giuridici.

Dal giro di orizzonte sulle tecniche giuridiche utili a regolare preventivamente il progresso tecnologico, emerge che il diritto possiede diverse alternative in grado di garantire la prevenzione dei danni, la previsione di alcuni rischi e la tu-

¹³ Cfr. *supra* nota 1.

tela dei diritti fondamentali degli utenti di tali tecnologie. È questa una precisazione doverosa, in quanto la questione della regolazione degli sviluppi tecnologici viene spesso accompagnata da una celata polemica relativa alla reale capacità del diritto di stare al passo con l'accelerazione tecnologica (*spacing problem*). L'atteggiamento polemico è proprio di certe correnti filosofiche, di cui diremo tra breve, ma non risparmia nemmeno i giuristi, che, tuttavia, sembrano non rendersi conto che, puntare troppo il dito sul c.d. *spacing problem*, rischia di fare il gioco di coloro che vogliono sganciare lo sviluppo tecnologico da ogni forma di vincolo regolativo. Chi assume questo atteggiamento non manca occasione per sottolineare che il diritto è inadatto a regolare il progresso scientifico e tecnologico perché intrinsecamente inadeguato a regolare ciò che è in divenire¹⁴.

Nel partito dei contrari all'intervento giuridico possiamo distinguere due principali tendenze. La prima è propria di coloro che ritengono che bastino codici di autocondotta da parte delle aziende per garantire un corretto uso della tecnologia. Si muovono in questa direzione le grandi multinazionali, che sposano l'ideologia transumanista. Questa posizione filosofica propone che sia il libero mercato a decidere le regole del gioco attraverso la legge della domanda e dell'offerta e vede nel diritto un ostacolo alla realizzazione di questo modello. In questa prospettiva, si chiedono ampi margini di manovra per le aziende e i finanziatori della ricerca, che, a loro volta, si traducono nel lasciare ampie zone grigie, ossia attività non necessariamente illecite, ma sostanzialmente fuori dal controllo del diritto statale e sovranazionale. Diversamente dai bioconservatori¹⁵, che, a loro volta, attraverso un uso ideologicamente orientato del diritto, cercano di mettere eccessivi lacci all'innovazione tecnologica, i transumanisti cercano di stare lontani dalle soluzioni giuridiche, almeno fintanto che gli interessi economici e di potere, relativamente a un dato settore, non siano riusciti a consolidarsi¹⁶. Delle tipologie di intervento giuridico richiamate in precedenza,

¹⁴Equilibrata e condivisibile l'opinione di Zaccaria che scrive «[i]n parallelo con il velocissimo sviluppo delle tecnologie bisogna rapidamente sviluppare e articolare nuovi strumenti concettuali e nuove categorie. Nonostante la sua nota capacità di rispondere in modo elastico ai mutamenti sociali, il diritto mostra di soffrire un certo ritardo dinnanzi alla complessità della materia e alla velocissima evoluzione della tecnologia», G. ZACCARIA, *Mutazioni del diritto: innovazione tecnologica e applicazioni predittive*, in *Ars Interpretandi*, 2021, 1, pp. 29-52, ivi, p. 31.

¹⁵T.L. EISSA, *Gesünder, intelligenter, perfekt?*, Freiburg-München, Verlag Karl Alber, 2014, come sottolinea l'autrice «[g]rundlegend für diese Positionen ist die Annahme, dass bestimmte menschliche Handlungen in sich und unabhängig von ihren Folgen verwerflich sind [...] Befürworter biokonservativen Denkens setzen darauf, die Endlichkeit und Unvollkommenheit des Menschen zu betonen und mit philosophischen und theologischen Argumenten in ihrer zentralen normativen Bedeutung aufzuzeigen», pp. 24-25. Si vedano, ad esempio, le posizioni di L.R. KASS, *Life, Liberty and the Defence of Dignity. The Challenge for Bioethics*, San Francisco, Encounter Books, 2002; M.J. SANDEL, *The Case against Perfection*, Harvard, Belknap Pr, 2009.

¹⁶Si ricorda la storia di Google raccontata da S. ZUBOFF, *Il capitalismo della sorveglianza. Il futuro dell'umanità nell'era dei nuovi poteri*, Roma, LUISS University Press, 2019.

vengono incentivate in quest'ottica le *regulatory sandboxes*, che consentono ampi spazi di manovra prima che le regole definitive siano elaborate.

Se quelli sinora richiamati sono alcuni dei principali aspetti problematici del rapporto tra tecnologia e diritto, ve ne è uno ulteriore che dipende dalla peculiarità delle tecnologie di cui trattiamo, ossia di essere considerate intelligenti. Sino ad ora abbiamo, infatti, analizzato la questione della tecnologia come oggetto di regolazione. In questa prospettiva, il diritto occupa il suo tradizionale posto e le traiettorie tecnologiche si muovono entro i confini di un ordinamento giuridico tradizionalmente inteso. Tuttavia, proprio la messa in campo di tecnologie intelligenti, in grado di incidere direttamente sul raggiungimento di obiettivi giuridici¹⁷, ha fatto emergere un diverso approccio. In questa inedita visione, la tecnologia non è oggetto di regolazione, bensì agente regolatore alla stregua del diritto.

2.2. Tecnologia come agente regolatore

Il cambio di paradigma suggerito con l'espressione "tecnologia come agente regolatore" è molto significativo e, al contempo, simbolico¹⁸. È incontestabile che, fino all'avvento delle tecnologie intelligenti, lo strumento più incisivo per controllare e guidare le condotte umane è stato il diritto, in virtù della sua forza coercitiva. Nell'era dell'IA ci si chiede se sia ancora così.

Di fatto, le possibilità offerte con le nuove tecnologie di agire direttamente sulle condotte individuali, ossia realizzando immediatamente l'effetto auspicato, rappresenta un'eccedenza su cui riflettere. Le applicazioni oggi disponibili per controllare da remoto strumenti e dispositivi, come ad esempio le automobili, consentono di agire direttamente sul dispositivo, rendendolo inutilizzabile dall'utente, di cui si vuole influenzare il comportamento. L'esempio, già richiamato nei capitoli che precedono, è quello dell'automobile che l'utente, inadempiente dal punto di vista assicurativo, non può più mettere in moto fino all'adempimento dell'obbligo. Altro esempio è la possibilità di limitare il numero di copie possibili di un'opera in formato digitale. In queste situazioni, lo strumento tecnologico consente di agire direttamente per ottenere il risultato voluto. Detto diversamente, in alcune circostanze la tecnologia determina direttamente ciò che si può o non si può fare. Occorre però intendersi sul significato da attribuire al

¹⁷ Si pensi alle tecnologie di profilazione usate in campo penale, o ancora alle tecnologie che consentono di incentivare il pagamento delle polizze assicurative dell'automobile, bloccando, in caso di ritardo, l'accensione del veicolo ecc.

¹⁸ J.R. REIDENBERG, *Lex Informatica: The Formulation of Information Policy Rules Through Technology*, in *Tex. L. Rev.*, 1998, 76, 3, pp. 553-593, ivi, p. 554, accessibile al link https://ir.law.net.fordham.edu/faculty_scholarship/42/; sul tema si veda anche L. LESSIG, *Code and other laws of cyberspace*, New York, Basic Books, 2009.

verbo “potere”. Si è osservato che con questi strumenti viene di fatto meno la possibilità di disobbedire alle norme, perché mentre le norme giuridiche «si limitano [...] a stabilire ciò che le persone *devono o non devono* fare, gli artefatti tecnologici hanno la capacità di determinare in maniera diretta ciò che le persone *possono o non possono* fare, rendendo superfluo l'intervento di una autorità di controllo che punisca, *ex post*, coloro che hanno violato la legge»¹⁹. In altre parole, la tecnologia rende impossibile agire diversamente, perché rende l'azione contraria all'obbligo difficile, se non impossibile, da realizzare. È l'azione che è resa impossibile di fatto. In tal senso, la tecnologia rende immediatamente esecutivo quello che, di norma, richiede un intermediario nel mondo giuridico (giudice).

La differenza che intercorre tra le norme di condotta di un sistema giuridico tradizionale e uno che può avvalersi di tecnologie, che consentono il controllo diretto sull'azione, sta nel modo in cui si guarda al “dovere” di fare qualcosa e al “potere” di farlo. Mentre le norme giuridiche, tradizionalmente intese, stabiliscono doveri di comportamento, il cui inadempimento richiede l'intervento di un terzo che, stabilito l'avvenuto inadempimento, dispone misure coercitive, la tecnologia intelligente consente, invece, di fare a meno di questo passaggio. Essa è in grado di rendere l'inadempimento di fatto impossibile²⁰. L'utilizzo delle tecnologie intelligenti, per ottenere questo preciso risultato, presenta certamente dei vantaggi in alcuni settori come quello contrattuale²¹ e quello assicurativo, dove l'intervento della tecnologia, direttamente inibitorio di certi comportamenti, può rappresentare una garanzia di certezza dell'esecuzione del contenuto contrattuale.

A fronte dei menzionati vantaggi, vanno, però, tenuti presenti anche alcuni aspetti problematici. In particolare, va evitato che tramite queste tecnologie, in grado di agire direttamente sull'azione da porre in essere, si instauri un'asimmetria tra i contraenti o si arrivi addirittura, in settori come quello delle condotte penalmente rilevanti, a manipolare il potere di scelta individuale, con interventi tecnologici che agiscono direttamente sul cervello. È già, infatti, possibile manipolare emozioni, memorie e sensazioni con l'ausilio delle neurotecnologie a scopo terapeutico – come vedremo nel dettaglio nel capitolo quarto. Nulla toglie che queste stesse tecnologie siano usate, un giorno non troppo lontano, per condizionare direttamente il comportamento di soggetti sani.

Va ulteriormente osservato che se attraverso la tecnologia è possibile agire immediatamente e direttamente sulla condotta dell'agente, da questa possibilità di intervento non è giocoforza dedurre che la tecnologia sia un agente regolato-

¹⁹N. LETTIERI, *Antigone e gli Algoritmi*, Modena, Mucchi, 2020, p. 32.

²⁰*Ibidem*.

²¹Cfr. B. CAPIELLO, G. CARULLO, *Gli smart contracts: prospettive per l'impiego nel settore pubblico-privato*, in R. CAVALLO PERIN, D.U. GALETTA (a cura di), *Il diritto dell'amministrazione pubblica digitale*, Torino, Giappichelli, 2020, pp. 305-320.

re, se con tale espressione vogliamo equipararla al diritto. L'uso dell'espressione "agente regolatore" è fuorviante sotto un duplice profilo. Innanzitutto, tale nozione, sfruttando la diffusa ideologia antropomorfizzante, promuove la falsa identificazione tra la tecnologia e quegli organi che, nei sistemi giuridici, sono deputati alla produzione normativa, come, ad esempio, il parlamento, o all'applicazione delle norme, come il giudice. Questa identificazione è perniciosa in quanto confonde le modalità coattive, che la tecnologia consente di mettere immediatamente in atto, con la spettanza del potere di decidere quali comportamenti regolare e con quali modalità. Se si tiene ferma questa distinzione ci si avvede che la tecnologia non è un agente regolatore, ma un possibile mezzo che si può decidere di adottare, in determinati settori e in relazione a specifiche attività, per ottenere certi risultati sul piano giuridico. In altre parole, è una decisione politica quella di lasciare che «contratti e leggi siano scritte direttamente nel codice del computer e siano suscettibili per ciò stesso di esecuzione automatica al verificarsi delle condizioni prestabilite – quanto il "techno-nudge", e cioè la promozione di comportamenti rispettosi delle norme attraverso una "spinta gentile" gestita autonomamente da un *software*»²². Se si accoglie questa prospettiva la tecnologia, lungi dal potersi qualificare come agente regolatore, risulta molto più umilmente mezzo di esecuzione delle norme. Questa scelta di delegare la modalità esecutiva alla tecnologia deve essere esplicitamente e chiaramente definita per evitare fraintendimenti sul ruolo degli attori in gioco, in una data relazione giuridica.

Essendo questa una scelta operata sul piano istituzionale e politico, occorre che siano chiari i contorni di tale scelta.

In breve, le scelte politico-istituzionali devono rimanere aderenti alla cornice dei diritti fondamentali. È questa la strada certamente perseguita in Europa nell'ambito della protezione dei dati. In questo contesto si è cercato di mantenere un equilibrio tra l'uso dei sistemi automatizzati a fini decisionali e il diritto di conoscere la logica di queste decisioni e la possibilità di richiedere l'intervento umano, laddove le decisioni automatizzate ricadano su sfere sensibili della vita degli individui.

Alla base della normativa sulla protezione dei dati vi è l'obiettivo di mantenere alto il controllo giuridico sul potere tecnologico. Attraverso questo controllo è possibile evitare derive nell'applicazione della *computational law*, mantenendo spazi in cui essa non trova applicazione. Dobbiamo interrogarci sull'opportunità di mantenere fermi ambiti di attività in cui la scelta di adempiere o meno a un dovere, previsto da una norma giuridica, segua il suo tradizionale percorso: dovere-inadempimento-intervento dell'autorità di controllo. Questa modalità regolativa si basa sull'idea che, conoscendo in anticipo le conseguenze

²² N. LETTIERI, *Antigone e gli Algoritmi*, cit. p. 34. Sul tema della tecnoregolazione anche U. RUFFOLO (a cura di), *Intelligenza artificiale – Il diritto, i diritti, l'etica*, Milano, Giuffrè, 2020.

dell'inadempimento, il soggetto resta libero di scegliere se aderire o meno alla prescrizione. La modalità richiamata si inserisce di norma in un modello normativo che Herbert Hart ha definito, a suo tempo, *choosing system*²³. È un modello che presuppone la libertà dei soggetti e la responsabilizzazione per le loro azioni. Se le norme giuridiche trovano attuazione mediante la tecnologia, che inibisce la possibilità di inadempimento in modo immediato, perché rende impossibile agire, il modello giuridico ispirato all'idea di un diritto come *sistema di scelta* viene di fatto ridimensionato, o addirittura superato, se gli ambiti, in cui si usano le tecnologie per impedire l'inadempimento, riguardano la grande maggioranza delle attività regolate dal diritto. A una superficiale analisi, avere a disposizione meccanismi che impediscono direttamente l'inadempimento di un obbligo può presentarsi come qualcosa di positivo e vantaggioso per le relazioni umane. Tuttavia, se si scava un po' più in profondità, ci si avvede di alcuni possibili svantaggi dell'applicazione generalizzata della tecnologia per finalità di controllo immediato dei comportamenti umani, vale a dire per impedire che si caratterizzino, attraverso l'inadempimento o l'agire contrario alla norma, come illeciti. Sembra un paradosso, ma non lo è. Se è vero che in settori come quello contrattualistico e assicurativo, l'uso a scopo coattivo della tecnologia presenta indubbi vantaggi in merito ai tempi dell'esecuzione e dell'efficacia del contratto, diversamente, in ambito penalistico, la riflessione non può fermarsi alla sola utilità pratica di questi dispositivi tecnologici. In questo specifico campo del diritto, risulta centrale la *vexata quaestio* della libertà di scelta come postulato dell'azione penale nei confronti di un individuo. In che senso? L'illecito penale ha, nell'immaginario collettivo, una valenza evocativa che non caratterizza illeciti di tipo civile o amministrativo. Ciò perché, sebbene gli illeciti penali presentino diversità sostanziali in termini di gravità, il diritto penale è storicamente, nella visione popolare, collegato alla repressione di illeciti che incidono in maniera profonda sulla vita e sulla libertà individuale. Se ciò è vero, è altresì vero che la storia del diritto penale è caratterizzata dal lento, ma incisivo, affrancamento dalla categoria dell'oggettività fino all'affermazione della responsabilità personale derivante da colpa legata alla volontà. Sebbene sussistano, anche negli ordinamenti moderni, forme di responsabilità oggettiva funzionali alla tutela di certi interessi, il postulato del diritto penale moderno è la libertà del soggetto di agire o di non agire in modo conforme alle norme e della conseguente responsabilità personale e colpevole. È sulla base di questo presupposto che si può ragionare in termini rieducativi e riabilitativi, anche attraverso percorsi di giustizia riparativa²⁴. A muovere da queste considerazioni, si può analizzare l'opportuni-

²³ È il modello di diritto «in which individuals can find out, in general terms at least, the costs they have to pay if they act in certain ways», H.L.A. HART, *Punishment and Responsibility. Essays in the Philosophy of Law*, Oxford, Clarendon Press, 1968, p. 44.

²⁴ Cfr. G. BERTAGNA; A. CERETTI, C. MAZZUCATO (a cura di), *Il libro dell'incontro*, Milano, Il Saggiatore, 2015.

tà di estendere o meno, a tutti gli ambiti del diritto, l'uso dell'IA, individuando appropriati criteri e modalità applicative. È sicuramente promettente l'impiego dell'IA ai fini di analisi e di interpretazione del diritto per creare tassonomie giuridiche basate sulla modellizzazione automatizzata di argomenti o per identificare e analizzare la giurisprudenza della Corte europea dei diritti dell'uomo su vari argomenti²⁵. Maggiore attenzione e cautela devono, invece, essere rivolte agli impieghi per l'applicazione dell'IA nel diritto penale. Questa attenzione va rivolta, innanzitutto, alle situazioni in cui decisioni automatizzate sono entrate nell'uso del sistema giudiziario. È questo il caso degli algoritmi attualmente in uso, ad esempio, negli Stati Uniti, per supportare i giudici nel valutare il rischio in relazione alle decisioni di messa in libertà provvisoria. Rispetto a questi algoritmi si pone un problema di eliminazione di *bias*, che possono condurre a decisioni discriminatorie, consolidando lo stigma sociale verso certe categorie di soggetti²⁶. Ma la riflessione va estesa, in prospettiva, a potenziali usi futuri. Vanno prese in considerazione quelle situazioni in cui un soggetto potrebbe, un giorno, tramite dispositivi collegati o inseriti nel corpo, essere direttamente controllato nelle azioni che compie ed essere immediatamente e direttamente disincentivato dal porre in essere un dato comportamento. Questo scenario chiama in causa le neurotecnologie che si avvalgono di *Brain-Computer Interfaces*. Attualmente queste tecnologie hanno prodotto importanti risultati in campo terapeutico, mostrando chiaramente la possibilità di intervenire sulla memoria e sulle emozioni²⁷. Non è, pertanto, fuori luogo pensare anticipatamente a loro eventuali impieghi al di fuori del contesto di cura e ponderare la loro applicabilità nei diversi ambiti di potenziale applicazione. L'uso delle neurotecnologie per controllare il comportamento dei consociati potrebbe essere di interesse pubblico quando applicato a coloro che, ad esempio, hanno già compiuto un reato di gravità medio-alta, per evitare la recidiva. Questo potenziale uso è quello che rischia di rimettere in discussione la libertà come postulato del diritto penale e indurre trasformazioni surrettizie del sistema penale moderno in sistema di *difesa e profilassi sociale*. Lo spostamento dal diritto penale come *sistema di scelta* al diritto penale come *difesa sociale* è di enorme portata, perché determina il passaggio da una visione dell'uomo libero e responsabile a quella di un soggetto determinato nel suo agire da forze da lui non controllabili.

In sintesi, alla base della prospettiva di un diritto penale principalmente, se non esclusivamente, orientato alla *difesa sociale*, vi è una visione etica senza libertà. L'adozione di tecnologie intelligenti in ambiti delicatissimi, in cui in gioco

²⁵ Cfr. F. DI PORTO, *La regolazione di fronte alle sfide dell'ICT e dell'Intelligenza Artificiale*, in R. CAVALLO PERIN, D.U. GALETTA (a cura di), Torino, Giappichelli, 2020, pp. 277-303.

²⁶ Cfr. R. LOWDEN, *Risk Assessment Algorithms: the Answer to an Inequitable Bail System?*, in *North Carolina Journal of Law & Technology*, 2018, 19, 4, pp. 221-251.

²⁷ Se ne parlerà in modo dettagliato nel capitolo quarto.

vi sono i valori di fondo su cui si basa la convivenza sociale, va valutata da chi è chiamato a decidere democraticamente e non lasciata a poteri occulti, che, per loro natura, garantiscono solo i propri interessi. Sono i parlamenti che devono decidere quale può essere l'impiego a fini coattivi della tecnologia. Se nulla toglie che l'innovazione tecnologica sia da considerarsi una risorsa preziosa anche per le attività del diritto, è altresì vero che anche in questo specifico contesto vanno evitati abusi che possono avere conseguenze anche molto devastanti, quando in gioco vi è la limitazione della libertà personale.

In breve, lungi dall'essere un agente regolatore al pari del diritto, la tecnologia è un mezzo al servizio degli obiettivi giuridici, che è responsabilità dei soggetti investiti della produzione normativa realizzare. Tenere ferma la distinzione tra ruolo del diritto e ruolo della tecnologia consente di evitare improprie invasioni di campo nella sfera del diritto, da parte di chi detiene il potere economico e tecnologico. Senza questa netta distinzione tra ciò che è puro mezzo (tecnologia) e ciò con cui si definiscono strategicamente i valori e i principi da perseguire (diritto inteso come mezzo con ruolo performativo), viene meno la possibilità di un dibattito chiaro sulle relazioni tra tecnologia e diritto e viene meno anche la possibilità di relazioni tra questi due mondi che siano funzionali ad entrambi.

3. Categorie giuridiche alla prova dell'IA

Le considerazioni affidate ai paragrafi precedenti sono state formulate con l'intento di evidenziare la complessità del rapporto tra scienza, tecnologia e diritto, soprattutto, in merito alla chiara definizione del ruolo di ciascuno di essi nel determinare le linee di azione e le direzioni dell'innovazione tecnologica, non meno che nell'identificazione dei soggetti destinati a operare tali scelte. Le scelte sui possibili usi delle tecnologie intelligenti non possono e non devono spettare esclusivamente ai tecnici, agli informatici o a chi ha interessi particolari di carattere economico. Le scelte spettano al decisore politico e a quei soggetti investiti della produzione normativa. Ad essi spetta il compito di stabilire scale di priorità e di valutare quali usi siano compatibili con i valori condivisi socialmente e culturalmente. Tuttavia, affinché il decisore politico possa prendere decisioni sostenibili sul lungo termine e congruenti con la cornice valoriale di riferimento, occorre che non si faccia tentare e influenzare da mode passeggero e da ideologie radicate nel senso comune, alimentate da dialettiche irrazionali. Mi riferisco all'ideologia antropomorfizzante e alla correlata dialettica degli opposti. Ci sono stati momenti della storia istituzionale europea in cui queste ultime sono entrate non solo nel dibattito politico, ma hanno rischiato di invadere anche il diritto e le sue categorie. Risale ormai al 2017 il dibattito, acceso e controverso, attorno alla Risoluzione del Parlamento europeo recante raccomandazioni

alla Commissione europea concernenti norme di diritto civile sulla robotica²⁸. La Risoluzione ha avuto un'importante eco mediatica²⁹, proprio in relazione alla ivi contenuta proposta di attribuire *e-personality* ai robot. In sostanza, si trattava di estendere l'applicabilità della categoria della personalità giuridica a robot dotati di specifiche caratteristiche: autonomia, capacità di autoapprendimento e imprevedibilità delle azioni³⁰. La proposta ha fatto molto discutere, tanto che molti esperti di diritto a livello internazionale si sono mobilitati sottoscrivendo una *lettera aperta*³¹ in cui chiedevano di non dare seguito alla richiesta del Parlamento. La Commissione europea nella sua risposta ufficiale del 2018 ha evitato di affrontare la questione, che è quindi oggi considerata superata. Se ciò è vero per il contesto europeo, altrettanto non si può dire per altre realtà in cui vi sono state scelte molto discusse di attribuire, ad esempio, cittadinanza a un robot³².

²⁸ Per un approfondimento cfr. S. SALARDI, M. SAPORITI, *Perché l'IA non deve diventare persona. Una critica all'ineluttabile 'divenire antropomorfo' delle macchine*, in ID. (a cura di), *Le tecnologie 'moralì' emergenti e le sfide etico-giuridiche delle nuove soggettività*, Torino, Giappichelli, 2020, pp. 52-74; S. SALARDI, *Robótica e inteligencia artificial: retos para el Derecho*, in *Derechos y Libertades*, 2020, 42, pp. 203-232.

²⁹ Cfr. S. MEDETTI, *E se l'Europa attribuisse una "personalità" elettronica ai robot?* In *Panorama*, 24 aprile 2018; L. FLORIDI, *L'ultima legge della robotica*, in *La Repubblica Robinson*, 12 febbraio 2017. Nello stesso periodo anche Bill Gates proponeva di tassare i robot, L. TREMOLADA, *Quella strana idea di tassare i robot*, in *Il Sole 24 Ore*, 21 febbraio 2017.

³⁰ Esistono una nozione ampia e una ristretta di robot. Cfr. S. NAVAS, *Robot Machines and Civil Liability*, in M. EBERS, S. NAVAS (eds.), *Algorithms and Law*, Cambridge, Cambridge University Press, 2020, pp. 157-173. La prima include macchine robotizzate come un braccio meccanico e anche entità intelligenti ed è la seguente «a system that is capable of perceiving the environment or context in which it is located, that can process the information to plan a certain action and execute it», ivi, p. 158. La seconda si riferisce solo a robot che hanno capacità di autoapprendimento e sono conosciuti come smart robot. Tale definizione esclude che macchine controllate da remoto dall'essere umano siano classificabili come robot, ivi, p. 160. Sul piano delle specifiche ricadute giuridiche si sostiene che «[a]i fini dell'ottimale imputazione delle responsabilità risarcitorie, l'attribuzione di personalità elettronica potrebbe [...] non essere la soluzione migliore. Essa, anzi, [...] potrebbe limitare e non accrescere le tutele dei soggetti lesi: i quali ad oggi godrebbero delle illimitate (e talora cumulative) responsabilità patrimoniali del fabbricante, o "custode", o utilizzatore, o "addestratore della "cosa" dotata di AI. Mentre la (sola) responsabilità diretta di quest'ultima se personificata, limiterebbe la responsabilità patrimoniale all'ammontare del patrimonio dell'ente robotico, così divenuto limitato capitale di rischio», cfr. U. RUFFOLO, *La "personalità elettronica"*, in ID. (a cura di), *Intelligenza artificiale – Il diritto, i diritti, l'etica*, 2020, Milano, Giuffrè, pp. 213-234, ivi, p. 216.

³¹ Cfr. il testo al link <http://www.robotics-openletter.eu>.

³² Al robot ginoide Sophia è stata attribuita la cittadinanza degli Emirati Arabi, <https://www.focus.it/tecnologia/innovazione/il-primo-robot-al-mondo-con-una-cittadinanza>. Sul tema cfr. F.G. PIZZETTI, *The Robot Sophia as a "New Citizen" of Saudi Arabia: what about granting legal personhood, "citizenship" and eventually dignity to non-human entities with artificial intelligence?*, in S. SALARDI, M. SAPORITI (forum a cura di), *Le tecnologie 'moralì' emergenti e i diritti fondamentali*

A ormai alcuni anni di distanza da quella discussa iniziativa contenuta nella richiamata Risoluzione, chi la ricorda in letteratura la definisce frutto di *sci-fi speculation* e si rallegra del fatto che nell'attuale proposta di Regolamento europeo non ve ne sia traccia³³. Se ciò può, in effetti, rasserenare gli animi, appare tuttavia opportuno ritornare su quella vicenda per almeno due ragioni. La prima ragione è far emergere quanto profondo sia il radicamento della visione antropomorfa della tecnica: si è arrivati, infatti, a toccare il cuore delle categorie giuridiche, quella di persona. La seconda ragione è comprendere le ragioni giuridiche della non estendibilità di questa categoria agli artefatti intelligenti. Per indagare tali ragioni occorre ricordare che il rapporto tra i giuristi e le categorie giuridiche si è basato, per un lungo periodo della storia del diritto, sulla concezione giusnaturalistica, che, negando la dimensione storica del diritto, ha incentivato una visione statica e fissa delle sue categorie. Il diritto inteso come diritto naturale è immutabile e mira a custodire i valori "perenni" dell'individuo. In tale prospettiva, le categorie giuridiche devono salvaguardare e realizzare un ordine predefinito e prestabilito delle cose al cui centro sta l'essere umano. Questi orientamenti giuridici sono stati fortemente influenzati dalla visione essenzialistica della realtà, a cui segue un'immagine unitaria, statica e immutabile dei concetti e delle categorie giuridiche.

Ma anche prescindendo dalla matrice giusnaturalistica, la fissità delle categorie giuridiche discende dall'attitudine concettualistica della dottrina giuridica, che ebbe, prima, nella *Scuola dell'Esegesi* francese e, successivamente, nella scuola tedesca della *Giurisprudenza dei concetti*, due fondamentali punti di riferimento. Sebbene questi orientamenti subirono profonde critiche³⁴, i loro assunti sono ancora latenti in atteggiamenti di molti giuristi moderni bioconservatori, che rifiutano il ripensamento delle categorie giuridiche, come operazione utile a far fronte a nuove e inedite situazioni. Nella prospettiva bioconservatrice, le categorie giuridiche non possono essere estese o interpretate diversamente da come

in Europa, in *Notizie di Politeia*, 2019, 133, pp. 63-70. Molto più recente è la notizia della possibile intestazione di un brevetto d'invenzione all'IA, presa in considerazione dall'ufficio dei brevetti del Sud Africa e dell'Australia, cfr. L. PEYRON, *Il caso del brevetto intestato a un robot e i limiti della tecnica*, in *Il Sole 24Ore*, 2 settembre 2021.

³³ Così, ad esempio, L. FLORIDI, *The European Legislation on AI: A Brief Analysis of its Philosophical Approach*, in *Philosophy and Technology*, 2021, 34, pp. 215-222, ivi, p. 218. C'è chi ricorda anche che la Risoluzione in questione «cited Mary Shelley's *Frankenstein* already at page 1, and invoked the creation of an Agency for the regulation of AI, as well as the attribution of legal personality, and thus "rights and duties" to smart autonomous robots» e qualifica questo inizio come *false start*, cfr. A. RENDA, *Moral Machines. The Emerging EU Policy on "Trustworthy AI"*, in W. BARFIELD (eds.), *The Cambridge Handbook of the Law of Algorithms*, Cambridge, Cambridge University Press, 2021, pp. 667-690, ivi, p. 667.

³⁴ Si pensi alla *Scuola del diritto libero* o alla *Giurisprudenza degli interessi* che fecero emergere la vaghezza, l'incoerenza, le lacune del diritto oltre a mettere in luce l'esistenza di altre fonti del diritto e la necessità di una certa discrezionalità per i giudici nella decisione delle controversie.

sinora fatto e, sulla base di questo assunto, si rifiuta l'estensione della personalità giuridica ai robot.

Tuttavia, questa giustificazione non appare sostenibile razionalmente. È stata la prospettiva filosofica di indirizzo analitico-linguistica a mettere in crisi la permanenza della convinzione relativa alla fissità delle categorie giuridiche. Nella prospettiva che guarda al diritto come linguaggio si è reso evidente che persona così come responsabilità e capacità sono concetti normativi, il cui significato è modulabile mediante lo strumento definitorio. A chi o cosa estendere la personalità giuridica, la capacità o la responsabilità è, innanzitutto, un'operazione di determinazione delle classi di cose o fatti denominati da una parola (denotazione o estensione) e dell'insieme delle proprietà che devono accomunare certe cose o certi fatti per appartenere alla classe denotata (connotazione o intensione). Dal punto di vista del significato giuridico dei termini, si tratta di una scelta interpretativa quella di denotare e connotare in maniera più o meno inclusiva una data parola. Ma proprio perché siamo nel mondo del linguaggio giuridico, la scelta interpretativa non è arbitraria, ma dipende da valutazioni di opportunità, unitamente a quelle di coerenza assiologica con i valori del sistema giuridico di riferimento. La scelta di estendere il significato di persona giuridica per includere i robot non è, pertanto, una scelta neutrale, scevra cioè da ricadute valoriali e simboliche. In altre parole, decidere di includere nel significato giuridico di persona i robot con certe caratteristiche non è un'operazione solo linguistica. Quando si individuano le proprietà rilevanti per stabilire e definire chi è persona giuridica e chi non lo è, si opera un bilanciamento di valori. Tale bilanciamento può ridurre o ampliare la categoria di soggetti o enti che hanno diritto di godere dei vantaggi di tale qualificazione giuridica. In questo senso, si spiegano le diversità storiche tra ordinamenti attorno a chi rientra o meno nella categoria di persona giuridica. Si pensi, ad esempio, alla posizione degli schiavi, che non sono persone; o alla posizione del nascituro, talora equiparato alla madre e tutelato come persona e talora tutelato come persona solo una volta nato; o ancora la scelta di alcuni ordinamenti contemporanei di attribuire i diritti alla natura e considerarla persona al pari dell'essere umano. È indubbio che l'estensione del pendolo semantico del concetto di persona porta seco molti vantaggi per gli enti inclusi. La posta in gioco è alta. Chi è persona sul piano giuridico gode della massima tutela da parte dell'ordinamento giuridico, in quanto destinatario dei diritti fondamentali. Quando l'estensione riguarda enti non umani – come ad esempio la natura, o nel caso da noi discusso di attribuzione di personalità giuridica ai robot – non si ha solo l'attribuzione formale di diritti e doveri, ma un'implicita equiparazione, dal carattere evocativo e simbolico, tra l'ente a cui è estesa la qualificazione giuridica e gli esseri umani. Chi ha optato per qualificare a livello costituzionale la natura come persona ha puntato proprio su questo importante effetto equiparativo. Lo stesso accadrebbe con i robot dotati di IA.

In sintesi, la questione di attribuire personalità giuridica ai robot mette in gioco diritti che si suole riferire agli esseri umani. Non vi sono ostacoli insuperabili sul piano della mera tecnica giuridica, perché persona non è una categoria fissa, ma un concetto normativo modulabile per gli scopi che il diritto si prefigge di raggiungere.

Quello che ci si deve chiedere, non è se l'operazione sia tecnicamente fattibile, perché lo è, ma se sia sostenibile eticamente. Per un'estensione della categoria della personalità giuridica ai robot occorre preliminarmente sciogliere i dubbi relativi: 1) ai motivi che spingono verso una tale scelta; 2) a chi è autorizzato a prendere una decisione così importante; 3) al destino della *persona costituzionalizzata*. La nozione di persona durante il processo di costituzionalizzazione ha perso, infatti, la sua accezione esclusivamente tecnica e neutrale, per divenire una nozione giuridica funzionale a costruire l'identità personale di ogni individuo e per identificare i valori che ispirano il sistema giuridico³⁵. Se questo è il confine, entro cui ci si muove nell'attuale momento storico, per definire chi è persona giuridica e chi non lo è, appare allora evidente che l'estensione della categoria per includere i robot intelligenti assume una valenza valoriale e simbolica, che va ben oltre l'utilità di trovare un regime giuridico adatto a risolvere le questioni relative alla responsabilità per danni. Questa operazione porta alla luce la profondità a cui l'ideologia antropomorfizzante si è insinuata anche nel pensiero istituzionale, che ne ha addirittura cercato una legittimazione giuridica. Eppure, sebbene i robot dotati di IA, di cui la citata Risoluzione tratta, abbiano particolari caratteristiche, sono pur sempre dispositivi «per cui è sufficiente prevedere una specifica e dettagliata regolamentazione della robotica (*robotics regulation*), muovendo dal presupposto che ad essere regolamentati non sono i robot in quanto tali, ma le condotte delle persone che interagiscono con loro, progettandoli, costruendoli, commercializzandoli, utilizzandoli»³⁶.

Richiamare la controversa vicenda istituzionale sull'*e-personality* non implica un giudizio negativo sull'innovazione tecnologica, vuole essere, invece, un modo per mettere in luce la strumentalizzazione di certe operazioni sul piano giuridico. Dietro tale proposta si cela il tentativo di rimbalzare le responsabilità per danni – causati da sistemi non facilmente controllabili – tra diversi attori: da chi li ha creati, a chi li ha programmati, a chi li ha prodotti fino all'utente finale.

La questione della responsabilità è stata affrontata nella recente Proposta di Regolamento europeo sull'IA³⁷. Diversamente dalla Risoluzione, la Propo-

³⁵ S. RODOTÀ, *Il diritto di avere diritti*, Roma-Bari, Laterza, 2012.

³⁶ S. AMATO, *Biodiritto 4.0. Intelligenza artificiale e nuove tecnologie*, Torino, Giappichelli, 2020, p. 91.

³⁷ Per un approfondimento sulla costruzione dell'etica dell'IA nel contesto europeo cfr. F.H. LLANO, *L'etica dell'intelligenza artificiale nel quadro giuridico dell'Unione europea*, in *Ragion Pratica*, 2021, 2, pp. 327-347.

sta³⁸ si concentra sul grado di pericolosità dei diversi sistemi di IA e adegua il sistema di responsabilità al rischio dei sistemi di IA e alla posizione dei diversi attori implicati nel processo di messa a disposizione e utilizzo dell'IA.

Si prevede una serie diversificata di obblighi in capo alle diverse figure (fornitori, fabbricanti, importatori, distributori, utenti), che si confrontano nei diversi stadi prima dell'utilizzo finale del prodotto e le cui responsabilità sono non solo diversamente distribuite in base al ruolo svolto nel processo, che va dalla progettazione allo sviluppo, alla immissione sul mercato, ma altresì in base al rischio che i sistemi di IA possono generare. L'approccio proporzionato basato sul rischio mira, infatti, a distinguere sistemi di IA a rischio inaccettabile e pertanto proibiti; ad alto rischio, ossia permessi, ma che devono essere adeguati ai requisiti richiesti dalla normativa (art. 16) e sottoposti a valutazione di conformità prima dell'immissione sul mercato o messa in servizio; a basso rischio, che richiedono l'adempimento di obblighi di trasparenza e informazione di cui all'articolo 52³⁹. L'approccio scelto risulta nel complesso condivisibile perché funzionale a mettere al riparo da iniziative esclusivamente pensate da soggetti privati attraverso l'autoregolamentazione. Quest'ultima rappresenta certamente un utile ausilio di *compliance* a norme stabilite dai decisori politici attraverso il meccanismo dell'incentivazione. Tuttavia, per evitare che l'autoregolamentazione divenga una scappatoia per imporre apparenti garanzie di fatto compatibili, più che altro, con gli interessi privati di chi le appresta, occorre che questi codici di autoregolamentazione vengano monitorati e debbano adeguarsi a standard previsti dalle norme generali del decisore politico. Si è già dimostrato ai tempi del dibattito sulla responsabilità penale degli enti come i codici di autoregolamentazione non rappresentino sempre l'effettiva e spesso occulta cultura di impresa⁴⁰, ma si possano configurare come una sorta di specchietto per le allodole.

³⁸Proposta di Regolamento del Parlamento e del Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale (Legge sull'intelligenza artificiale) e modifica alcuni atti legislativi dell'Unione, COM (2021) 206 final. Ma il tema della fiducia è presente anche in altre strategie come quella statunitense, *Final Report National Security Commission on Artificial Intelligence*, 2021, <https://www.nscat.gov/wp-content/uploads/2021/03/Full-Report-Digital-1.pdf>.

³⁹L'articolo 52 della Proposta di Regolamento recita al comma 1 che «i fornitori garantiscono che i sistemi di IA destinati a interagire con le persone fisiche siano progettati e sviluppati in modo tale che le persone fisiche siano informate del fatto di stare interagendo con un sistema di IA, a meno che non risulti evidente dalle circostanze e dal contesto di utilizzo». La bozza è attualmente in discussione e molti sono gli emendamenti presentati.

⁴⁰Nel tentativo di trovare strade giuridiche per responsabilizzare gli enti e le persone giuridiche è stata elaborata la c.d. colpevolezza come manifestazione della cultura di impresa, che ha trovato esplicita elaborazione nel *Criminal Code Act* australiano del 1995. Lo scopo di tale istituto è quello di incidere non solo sulle forme ufficiali di comportamenti illeciti degli enti, ma di poter raggiungere situazioni «in cui al dispetto dell'esistenza di documenti formali che proclamano il

In relazione alla richiamata proposta di Regolamento europeo sull'IA, pare opportuno svolgere qualche ulteriore riflessione sulla scelta linguistica di espressioni come "ecosistema di fiducia", "IA affidabile" (*trustworthy*)⁴¹ e in genere sulla decisione di collegare il rapporto che l'utente sviluppa nei confronti dei prodotti dotati di IA alla nozione di fiducia. L'intento di questa riflessione sulla fiducia è di completare l'analisi effettuata nel primo capitolo su tale nozione, con uno sguardo specificamente rivolto, in questa sede, al diritto.

Incentrare il rapporto sulla fiducia è piuttosto fuorviante, ma sicuramente funzionale al marketing dei prodotti⁴².

Il focus sulla fiducia inibisce, di fatto, l'attenzione per le questioni etiche e riduce il ruolo dell'utente finale a quello di mero utilizzatore passivo di quanto proposto. Egli deve *fidarsi* del fatto che i prodotti che rispettano certi standard sono utilizzabili in sicurezza. La fiducia e il suo ecosistema si fondano, quindi, sulla prospettazione di corrette soluzioni tecniche. In poche parole, nella Proposta le soluzioni tecniche sono la base della fiducia e non vi è necessità di prospettare altri scenari. Questo è il messaggio di base che viene recepito nella *folk psychology* e che è funzionale alla promozione pubblicitaria dell'IA.

Questa impressione è supportata dalle dichiarazioni di Thomas Metzinger, membro del gruppo di esperti che ha lavorato al documento, uscito nel 2019, *Orientamenti etici per un'IA affidabile*⁴³. Il professore così si esprime «[...] the Trustworthy AI narrative is, in reality, about developing future markets and using ethics debates as elegant public decorations for a large-scale investment strategy. At least that's the impression I am beginning to get after nine months of working on the guidelines»⁴⁴.

Non poche sono le insidie della scelta semantica e politica di spostare l'at-

rispetto della legge, ci si aspetta in realtà una violazione delle regole», cfr. C. DE MAGLIE, *L'etica e il mercato*, Milano, Giuffrè, 2002, p. 359.

⁴¹ Tra le varie *milestones* indicate sul sito dell'Unione europea in tema di IA risultano particolarmente esemplificativi i titoli utilizzati. Già nel 2019 la questione etica è inesorabilmente associata alla questione della fiducia con le *Ethics Guidelines for Trustworthy AI*, ma da lì in poi nei titoli appare solo la fiducia: 2019 *Communication: Building Trust in Human Centric Artificial Intelligence*; 2019 *Piloting of Assessment list on Trustworthy AI*; 2020 *White Paper on AI: A European approach to excellence and trust*, cfr. il link <https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence>.

⁴² M. CHIRIATTI, *Incoscienza artificiale. Come fanno le macchine a prevedere per noi*, Roma, LUISS University Press, 2021, p. 47.

⁴³ Reperibile al link <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>.

⁴⁴ T. METZINGER, *EU Guidelines. Ethics Washing Made in Europe*, in *Tagesspiegel*, 8 aprile 2019, accessibile al link <https://www.tagesspiegel.de/politik/eu-guidelines-ethics-washing-made-in-europe/24195496.html>. Tuttavia la nozione di *ethics washing* è stata criticata dal *Panel for the Future of Science and Technology* nel suo studio del 2020 intitolato *Artificial Intelligence: from ethics to policy*, p. 3, testo accessibile al link [https://www.europarl.europa.eu/RegData/etudes/STUD/2020/641507/EPRS_STU\(2020\)641507_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2020/641507/EPRS_STU(2020)641507_EN.pdf).

tenzione dall'etica alla fiducia. Il concetto di fiducia tende a offuscare il richiamo all'autonomia decisionale per esaltare, invece, l'aspetto dell'affidamento. È vero che è una fiducia prestata sulla base di garanzie tecniche, come elaborate dalla Proposta di Regolamento.

Tuttavia, è una fiducia quasi cieca perché l'utente medio non è in grado di fare alcuna valutazione sull'affidabilità tecnica del prodotto che usa e, per tale motivo, nemmeno sollevare eventuali dubbi sull'opportunità di utilizzare proprio quel prodotto. Per certi versi, siamo in una situazione in cui si chiede agli utilizzatori di affidarsi, senza chiarire chi è davvero il destinatario di quella fiducia: l'IA, chi programma, chi produce, chi commercializza queste tecnologie?

La strategia che fa leva sull'affidamento dell'utente appare la più veloce e meno onerosa rispetto a quelle strategie che richiedono una seria alfabetizzazione degli utenti, intesa non solo come sviluppo di competenze tecniche, ad esempio di informatica per l'utilizzo dei servizi digitali⁴⁵, ma come ampio lavoro educativo, volto alla sollecitazione di un approccio critico.

Concludendo, affinché la narrazione della *trustworthy AI*, che nella prospettiva europea dovrebbe porre le basi per una *human-centric vision* del progresso tecnologico, non sia meramente una formula retorica, occorre che la costruzione dell'ecosistema di fiducia parta dalla promozione dell'alfabetizzazione seria degli utenti. Se fondata sull'autonomia individuale, la fiducia non è più cieca, ma si configura come un processo in divenire di costruzione della relazione con il macchinico, che non si ritiene mai completamente compiuto. Al centro di tale processo c'è sempre la persona umana come misura della fiducia.

4. I diritti fondamentali sono la cornice adeguata a regolare l'innovazione tecnologica? Il caso del riconoscimento facciale

La domanda posta nel titolo di questo paragrafo ci spinge a riflettere sull'adeguatezza dell'attuale cornice europea dei diritti fondamentali⁴⁶ per orientare

⁴⁵ Vedasi articolo 8 del Codice dell'amministrazione digitale, decreto legislativo 7 marzo 2005, n. 82. Cfr. anche EUROPEAN COMMISSION, *The European Digital Competence Framework for Citizens*, 2016 reperibile a questo link [file:///Users/silviasa/Downloads/KE-02-15-657-EN-N%20\(2\).pdf](file:///Users/silviasa/Downloads/KE-02-15-657-EN-N%20(2).pdf). Guido Scorza, componente del Garante della Privacy parla dei problemi di alfabetizzazione della popolazione italiana in relazione agli *smart glasses* in un'intervista dal titolo *Smart glasses, il ritorno. Ma c'è un problema di privacy* del 20 settembre 2021, disponibile al link <https://www.garanteprivacy.it/home/docweb/-/docweb-display/docweb/9701499>.

⁴⁶ Le nostre riflessioni saranno circoscritte al contesto europeo e alla traduzione dei diritti fondamentali in questo specifico ambito geografico, sebbene il linguaggio dei diritti umani sia da tempo un linguaggio internazionale. Tale scelta ci consente di fare affermazioni sui diritti che hanno una maggiore solidità e coerenza e che possono contare, nel contesto europeo, su una convergenza circa i valori di cui sono espressione, rispetto alla maggiore fluidità che caratterizza l'ar-

l'innovazione tecnologica. La pervasività di questa innovazione richiede un'attenta ponderazione dei rischi e dei benefici e sono i diritti fondamentali a rappresentare la misura di questa ponderazione.

L'appello ai diritti costituisce ad oggi la strategia politica utilizzata dalle istituzioni europee in tema di regolazione dell'IA. La stessa Proposta di Regolamento richiama l'interesse ad «assicurare che i cittadini europei possano beneficiare di nuove tecnologie sviluppate e operanti in conformità ai valori, ai diritti fondamentali e ai principi dell'Unione»⁴⁷.

Si pone, pertanto, il problema della garanzia dell'effettività dei diritti fondamentali, che non è un problema nuovo. Norberto Bobbio aveva posto questa questione nel 1996, nel suo volume intitolato *L'età dei diritti*⁴⁸. Come osserva Bobbio, non si tratta tanto di individuare nuovi diritti, quanto di garantire efficacemente la tutela di quelli esistenti. In altre parole, non basta avere sulla carta i diritti, ma vanno portati il più possibile a compimento. È questo, però, l'aspetto più difficoltoso da realizzare e che trova ostacoli anche nella *culla dei diritti fondamentali*. La tensione tra diritti fondamentali, intesi come strategia di tutela del soggetto più debole di una qualunque relazione giuridica, e l'aspirazione a chi detiene un qualunque tipo di potere, da quello statale a quello economico e tecnologico, di esercitarlo *legibus solutus* è, ad oggi, una questione irrisolta. I poteri forti mal si adattano ai limiti imposti tramite i diritti fondamentali e tentano nuove strade per aggirarli o addirittura depotenziarli. La strategia non consiste di norma nel riconoscersi in modo esplicito «nemico dei diritti», ma nel promuovere *atteggiamenti ipocriti e farisaici* rispetto ai diritti⁴⁹, come l'adozione di formule compromissorie nell'elaborazione delle norme.

gomento dei diritti umani a livello internazionale, ad esempio intorno alla loro universalità. Nel contesto europeo, infatti, la spinta verso la positivizzazione dei diritti e il superamento in tal modo di alcune questioni ancora aperte sul piano internazionale è, da tempo, realizzata. Possiamo sostenere che la definizione di diritti come diritti fondamentali nel contesto europeo si configura come una questione strutturale, cfr. U. SCARPELLI, *Diritti positivi, diritti naturali: un'analisi semiotica*, in S. CAPRIOLI, F. TREGGIARI (a cura di), *Diritti umani e civiltà giuridica*, Perugia, Centro Studi Giuridici e Politici, 1992, pp. 31-44. Inoltre, se le nostre riflessioni si riferiscono al contesto europeo, ciò consente anche di non dover necessariamente trattare in questa sede l'annosa questione del c.d. eurocentrismo e del controverso legame dei diritti fondamentali con la cultura europea, cfr. A. FACCHI, *Breve storia dei diritti umani*, Bologna, Il Mulino, 2007, p. 146. Circoscrivere al contesto europeo la discussione aiuta anche a superare l'antitesi tra diritti fondamentali come diritti morali generali e diritti giuridici, che alcuni autori propongono, cfr. A. BUCHANAN, *The Heart of Human Rights*, New York, Oxford University Press, 2013. Sul tema dei diritti umani cfr. anche M. SAPORITI, *Dialettiche sovrane. Uno studio sul rapporto tra sovranità e religioni*, Torino, Giapichelli, 2019, in particolare il terzo capitolo.

⁴⁷ Così la relazione introduttiva alla Proposta di Regolamento europeo sull'IA.

⁴⁸ N. BOBBIO, *L'età dei diritti*, Torino, Einaudi, 1997.

⁴⁹ Mutuando le riflessioni di F.J. ANSUÁTEGUI ROIG, *Norme, Giudici, Stato Costituzionale. Frammenti di un Modello Giuridico*, cit. p. 25, possiamo guardare ai diritti fondamentali «come una strategia umana attraverso la quale si concretizza l'impegno a favore della tutela delle libertà».

Queste formule non risultano tanto incoerenti sul piano di tecnica del diritto, quanto assiologicamente incoerenti con la visione valoriale espressa dalla cornice dei diritti. Un esempio è dato dalla previsione inserita nell'attuale Proposta di Regolamento europeo sull'IA sul riconoscimento facciale. Come osservato in precedenza, l'approccio della proposta si basa sul rischio. A tal fine sono state previste specifiche categorie di tecnologie che sono catalogate secondo una logica piramidale di decrescente (in)accettabilità: si va dal rischio inaccettabile con conseguente divieto di certe pratiche di intelligenza artificiale, previste dall'articolo 5; ai sistemi ad alto rischio, disciplinati dall'articolo 6 e seguenti; ai sistemi a basso rischio, regolati dall'articolo 52, per i quali sono previsti solo specifici obblighi di trasparenza.

Il riconoscimento facciale, di cui parliamo, ricade, nella proposta di Regolamento, sotto la categoria dei "sistemi di identificazione biometrica⁵⁰ remota in tempo reale in spazi accessibili al pubblico". Si tratta di quelle forme di identificazione biometrica in cui il rilevamento dei dati biometrici, il confronto e l'identificazione avvengono senza ritardi significativi, in luoghi fisici accessibili al pubblico, indipendentemente dall'applicabilità di determinate condizioni di accesso⁵¹. L'identificazione biometrica remota in tempo reale in spazi accessibili al pubblico rientra tra quelle pratiche di IA vietate dall'articolo 5. Tuttavia, non si

In quanto strategia umana, come ogni strategia giuridica, politica e morale, siamo noi i protagonisti e i responsabili del suo successo o del suo fallimento». Anche Alessandra Facchi ricorda che «nel dopoguerra i diritti dell'uomo tornano a essere una categoria a cui affidare la funzione fondamentale di proteggere la libertà degli individui, anche di quelli che non si rispecchiano nella maggioranza, determinando limiti e impegni per i governi», A. FACCHI, *Breve storia dei diritti umani*, cit. p. 133.

⁵⁰Per sistemi di identificazione biometrica si intende «[b]iometrics technologies are used to identify, verify, or confirm a person's identity based on their physiological (external appearance) or behavioural (how they act) characteristics. Physiological characteristics are assessed through morphological identifiers (mainly consisting of fingerprints, the hand's shape, the finger, vein pattern, the eye (iris and retina), and the face's shape) and biological analyses (DNA, blood, saliva, or urine). Behavioural characteristics are commonly assessed using voice recognition, signature dynamics (speed of movement of pen, accelerations, pressure exerted, inclination), gait (i.e. individual walking style) or gestures. Biometrics allows a person to be identified and authenticated based on verifiable unique and specific data. Biometric identification consists of determining the identity of a person by capturing an item of their biometric data (e.g. a photograph) and comparing it to the biometric data of several other persons kept in a database, providing an answer to the question "Who are you?". Biometric authentication compares data on a person's characteristics to their biometric data to determine resemblance and provides an answer to the question "Are you Mrs or Mr X?". Biometric technologies include "fingerprint recognition", "signature recognition", "DNA matching", "eyes – iris recognition", "eyes – retina recognition", "voice – speaker identification", "gait", "hand geometry recognition" or "face recognition"», cfr. EUROPEAN PARLIAMENT, *Regulating Facial Recognition in the EU*, 2021, elaborato da l'European Parliamentary Research Service, accessibile al link [https://www.europarl.europa.eu/RegData/etudes/IDAN/2021/698021/EPRS_IDA\(2021\)698021_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/IDAN/2021/698021/EPRS_IDA(2021)698021_EN.pdf).

⁵¹Cfr. articolo 3 della Proposta di Regolamento.

tratta di un divieto assoluto, come quello previsto per il *social scoring*⁵². È un divieto che riguarda solo le attività di contrasto da parte delle forze dell'ordine, per le quali sono previste alcune eccezioni quando tali attività perseguono i seguenti obiettivi: 1) la ricerca mirata di potenziali vittime specifiche di reato, compresi i minori scomparsi; 2) la prevenzione di minaccia specifica, sostanziale e imminente per la vita o l'incolumità fisica di persone fisiche o di un attacco terroristico; 3) il rilevamento, la localizzazione, l'identificazione o l'azione penale nei confronti di un autore o un sospettato di reato di cui all'articolo 2, paragrafo 2, della decisione quadro 2002/584/GAI, del Consiglio che è relativa al mandato di arresto europeo e alle procedure di consegna tra Stati membri.

Un primo aspetto spinoso è che questi sistemi sono vietati solo alle forze dell'ordine, mentre enti privati che soddisfano certi requisiti, stabiliti dalla proposta di Regolamento, potrebbero utilizzarli. Questo comporta il rischio di un riconoscimento facciale di massa⁵³. Di fatto, anche se sono previsti criteri di controllo come la procedura di valutazione della conformità dei sistemi usati, risulterà molto difficile garantire un'effettiva tutela dei diritti che, con queste pratiche di IA, possono essere violati. Non si tratta solo della privacy, che è certamente il nodo centrale del *bundle of rights* che viene messo in pericolo. A seguito della violazione della privacy, si possono verificare ulteriori discriminazioni, anche gravi, nei confronti dei soggetti più vulnerabili. In gioco vi sono, pertanto, la libertà personale e il correlato diritto al libero sviluppo della propria personalità, che la riduzione, se non addirittura l'eliminazione, di spazi dove circolare anonimamente, mette in grave crisi. Un massiccio impiego del riconoscimento facciale può indurre il conformismo dei comportamenti, che non va confuso con

⁵² Trattasi di quelle pratiche disciplinate dall'articolo 5 comma 1 lettera c, utilizzate da autorità pubbliche o per loro conto ai fini della valutazione o della classificazione dell'affidabilità delle persone fisiche per un determinato periodo di tempo sulla base del loro comportamento sociale o di caratteristiche personali o della personalità note o previste, in cui il punteggio sociale così ottenuto comporti il verificarsi di uno o di entrambi i seguenti scenari: 1) un trattamento pregiudizievole o sfavorevole di determinate persone fisiche o di interi gruppi di persone fisiche in contesti sociali che non sono collegati ai contesti in cui i dati sono stati originariamente generati o raccolti; 2) un trattamento pregiudizievole o sfavorevole di determinate persone fisiche o di interi gruppi di persone fisiche che sia ingiustificato o sproporzionato rispetto al loro comportamento sociale o alla sua gravità. Cfr. N.G. PACKIN, Y. LEV-ARETZ, *Algorithmic Analysis of Social Behaviour for Profiling, Ranking, and Assessment*, in W. BARFIELD (eds.), *The Cambridge Handbook of The Law of Algorithms*, 2021, pp. 632-663. Secondo gli autori bisogna stare molto attenti a non incorrere nell'errore di ritenere il modello cinese molto distante dal contesto occidentale «[a]nd while to people who live in Western democracies the Chinese model feels like a fat-fetched dystopia, evidence shows that scoring systems developed by powerful private players can, in fact, be even worse», ivi, p. 653. Motivo in più per ritenere che questa pratica debba rimanere vietata, perché palesemente contraddice il modello democratico fondato sui diritti fondamentali presente in Europa.

⁵³ Cfr. *Europe's AI rules open door to mass use of facial recognition, critics warn*, articolo di Melissa HEIKKILÄ in *Politico*, 7 giugno 2021.

la scelta consapevole di tenere comportamenti conformi alle regole di condotta. È un conformismo inteso come omologazione, che cancella la diversità di espressione individuale.

Un secondo insidioso problema della norma giuridica contenuta nella Proposta sul tema del riconoscimento facciale è che rappresenta una possibile strategia di svuotamento dei diritti e della loro funzione di *empowerment of ordinary people*.

Va, infatti, considerato che, quando è stata presentata la Proposta, diversi attori – esponenti politici, membri del gruppo di esperti che hanno lavorato alle linee guida europee per l’etica dell’IA uscite nel 2019, appartenenti alla società civile – si sono chiaramente espressi contro il compromesso sul riconoscimento facciale raggiunto e poi confluito nella Proposta. Queste voci contrarie hanno trovato espressione anche nell’analisi dell’*European Parliamentary Research Service* dedicata al tema del riconoscimento facciale⁵⁴. A livello nazionale, lo stesso Garante per la Protezione dei Dati Personali ha espresso, in linea con il Consiglio d’Europa, parere contrario all’uso di un particolare sistema di riconoscimento facciale *Sari Real Time* nel suo parere del 25 marzo 2021⁵⁵. Tra le critiche mosse a un riconoscimento facciale esercitato in maniera indiscriminata vi è quella cruciale che sottolinea come tale pratica inibisca il libero sviluppo della personalità.

Con riguardo a quest’ultimo diritto, trattasi di un diritto centrale del costituzionalismo europeo, che ha consentito di superare l’astrazione del soggetto giuridico per calarlo dentro l’esperienza concreta, nella quale deve trovare forme di realizzazione personale, congruenti con la sua specificità. In questa prospettiva, l’individuo diviene «costruttore della propria personalità»⁵⁶. Affinché questa autorealizzazione possa trovare spazio in una società permeata dalla tecnologia, occorre chiedersi se, per alcuni usi, non sia opportuno tracciare linee rosse convincenti, come potrebbe essere il caso di divieto per il riconoscimento facciale. In assenza di una presa di posizione decisa, si rischia che i compromessi, palesemente insostenibili sul piano valoriale rispetto ai contenuti dei diritti fondamentali, compromettano la realizzazione di quell’ecosistema di fiducia, a cui la Proposta sembra mirare.

⁵⁴ EUROPEAN PARLIAMENT, *Regulating Facial Recognition in the EU*, cit. «While stakeholders, researchers and regulators seem to agree on a need for regulation, some critics question the proposed distinction between low-risk and high-risk biometric systems and warn that the proposed legislation would enable a system of standardisation and self-regulation without proper public oversight. They call for amendments to the draft text, including with regard to the Member States’ leeway in implementing the new rules. Some strongly support stricter rules – including an outright ban on such technologies» p. I.

⁵⁵ Il parere del Garante è reperibile al link <https://www.garanteprivacy.it/web/guest/home/docweb/-/docweb-display/docweb/9575877>.

⁵⁶ S. RODOTÀ, *Il diritto di avere diritti*, cit. p. 263.

La cornice dei diritti fondamentali consente l'autorealizzazione solo se presa sul serio e portata a compimento. La sua adeguatezza a tenere il passo con le profonde trasformazioni dell'umano, indotte dalla scienza e dalla tecnologia, dipende, innanzitutto, dalla serietà con cui si affronta la questione dei valori che debbono prevalere nel bilanciamento dei vari interessi. Il principio del libero sviluppo della propria personalità rappresenta la stella polare per concretizzare il generale e costante richiamo al fatto che si deve *mantenere l'uomo al centro*⁵⁷. Quest'espressione è piuttosto generica in questa formulazione e affinché, a sua volta, non sia solo un bello slogan, va resa effettiva.

4.1. *Il principio di precauzione e la tutela effettiva dei diritti*

Con riferimento alle tecnologie intelligenti, di cui discutiamo in questo volume, mantenere la centralità del principio di precauzione appare funzionale all'obiettivo di autorealizzazione dell'individuo. Questo principio, ideato proprio per far fronte alle incertezze profonde proiettate nel futuro dello sviluppo scientifico e tecnologico, è da tempo un principio ispiratore delle politiche europee in materia ambientale, ma estende il suo raggio di azione ben oltre quell'ambito. È un principio che ispira il processo decisionale quando l'incertezza sui rischi è alta e per lo più sconosciuta. In senso lato, si può leggere il principio di precauzione come una richiesta di cautela e previdenza quando si tratta di prendere decisioni che concernono attività con ricadute potenzialmente negative e ignote circa la loro incidenza. In un'accezione più ristretta, tale principio richiede che si valuti un intervento normativo, eventualmente anche proibitivo se necessario, anche se non sussiste prova certa del potenziale danno. Si tratta di un principio che ha una capacità riconciliatrice tra diversi interessi in gioco, ma che tende a tutelare chi è nella condizione di maggiore debolezza e come tale può considerarsi «come metodo al quale ispirare i necessari interventi di una regola giuridica capace di mantenere fermo il riferimento ai diritti fondamentali»⁵⁸.

Tale principio è più incisivo rispetto al principio di prevenzione, dal quale si differenzia sottilmente, ma sostanzialmente, perché consente di non fermarsi a rischi oggettivi e provati, ma di decidere di considerare l'eccessiva incertezza per

⁵⁷ Sembra andare in questa direzione la proposta da parte della Commissione europea di una *Dichiarazione sui diritti e principi digitali per tutti nell'UE*, dove si affermano: il primato della persona nel processo di trasformazione digitale; la rilevanza della solidarietà e dell'inclusione; la libertà di scelta e la conseguente trasparenza degli algoritmi; la partecipazione allo spazio pubblico digitale; la sicurezza online; economia circolare dei prodotti digitali.

Cfr. file:///Users/silviasa/Downloads/La_Commissione_propone_una_dichiarazione_su_diritti_e_principi_digitali_per_tutti_nell_UE.pdf; e l'articolo pubblicato sul Wired reperibile a questo link <https://www.wired.it/article/europa-diritti-digitali/>.

⁵⁸ S. RODOTÀ, *Il diritto di avere diritti*, cit. p. 371.

prevedere misure incisive di tutela dei diritti fondamentali. Il principio di precauzione rappresenta un utile mezzo di orientamento nelle decisioni sull'innovazione tecnologica, se non è utilizzato in maniera strumentale per difendere posizioni ideologicamente contrarie allo sviluppo scientifico e tecnologico. Non è un principio contro la scienza o la tecnologia, bensì contro lo strapotere e l'abuso dell'innovazione scientifico-tecnologica e contro l'inedita eccedenza di cui tale potere dispone. In altre parole, si pone come argine contro strategie di svuotamento dei diritti fondamentali. Ecco perché appare criticabile che la Proposta di Regolamento europeo sull'IA non ne faccia menzione esplicita. È auspicabile che sul punto si torni a riflettere, vista anche la valenza morale sottesa a tale principio.

Tale valenza morale si esprime attraverso il riconoscimento del fatto che ciò che crea l'uomo può produrre rischi non identificabili a priori e porta con sé l'umile accettazione della *difettività* umana nella sua capacità creatrice. Riconoscere questa difettività non significa necessariamente condannarla, ma trovare strade per tutelare l'umano e fare di questa *difettività un parametro del governo della vita pubblica*⁵⁹. In questa prospettiva, sarà più agevole contrastare l'ingenuo, ma insidioso, antropomorfismo⁶⁰ del macchinico, di cui si è dato conto in questo volume.

Un ulteriore sforzo, richiesto ai decisori politici, che prendono i diritti fondamentali sul serio, è quello di mantenere fermo e attuare il principio di esplicabilità (*explicability*).

4.2. Il principio di esplicabilità e i diritti portati a compimento

Questo principio compare in maniera esplicita nel 2018 nel documento *AI4People's Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations*⁶¹. Il documento mutua dal contesto bioetico quattro principi per orientare l'innovazione tecnologica, ossia beneficenza, non-maleficenza, giustizia e autonomia⁶². Ai redattori del documento apparve op-

⁵⁹ A. PUNZI, *Difettività e giustizia aumentata. L'esperienza giuridica e la sfida dell'umanesimo digitale*, in *Ars Interpretandi*, 2021, 1, pp. 113-128, ivi, p. 120.

⁶⁰ S. RODOTÀ, *Il diritto di avere diritti*, cit. p. 345.

⁶¹ Il documento riassume gli esiti dell'iniziativa AI4PEOPLE lanciata in quell'anno da ATOMIUM-EISMED, <https://www.eismd.eu/>.

⁶² Per ciascuno di questi principi si individuano termini esplicativi. Per il principio di beneficenza si richiamano *promoting well-being, preserving dignity, and sustaining the planet*; per il principio di non-maleficenza si fa riferimento a *privacy, security and "capability caution"*; per il principio di giustizia si identificano *promoting prosperity and preserving solidarity*; per il principio di autonomia si sottolinea *the power to decide (whether to decide)*. Per un approfondimento si rimanda al documento accessibile al link <https://www.eismd.eu/wp-content/uploads/2019/03/AI4PEOPLE%E2%80%99s-Ethical-Framework-for-a-Good-AI-Society.pdf>.

portuno aggiungerne un quinto principio, specificamente dedicato alle peculiarità delle tecnologie intelligenti, ossia l'*explicability*, che si ritiene sia il mezzo per *enabling the other principles through intelligibility and accountability*.

Questo principio è utile ad affrontare due questioni: una tecnica, che riguarda la possibilità di comprendere come un sistema artificiale produca certi *outputs*; una etico-giuridica concernente la responsabilità, o meglio l'identificazione del soggetto responsabile del (mal)funzionamento dell'algoritmo. Questo principio si ricollega, quindi, alle problematiche tecniche dell'opacità di alcuni sistemi algoritmici c.d. *black box*, di cui abbiamo dato conto nel capitolo precedente. Tale principio è stato nel tempo ripreso diffusamente non solo in letteratura⁶³, ma anche a livello istituzionale europeo⁶⁴ attraverso la correlata nozione di trasparenza⁶⁵. Non è intenzione di chi scrive né di ripercorrere le problematiche tecniche che attengono al raggiungimento di una *esplicabilità* completa allo stato attuale dello sviluppo tecnologico, che per l'appunto potrà trovare soluzioni innovative con il passare del tempo, né soffermarsi sull'analisi delle varie nozioni correlate al principio richiamato. Quello che si intende fare è spendere qualche argomento a favore del mantenimento di questo principio, nei limiti della sua possibile attuazione tecnica, come punto fermo per la realizzazione dell'IA c.d. *human-centric* e della correlata attuazione del diritto al libero sviluppo della propria personalità.

Vi sono attività umane rispetto alle quali l'irrinunciabilità a questo principio e allo sforzo tecnico per garantire l'esplicabilità e la comprensibilità dei sistemi intelligenti, al più alto grado possibile, dovrebbe essere condivisa da tutti coloro che dicono di avere a cuore il futuro dell'umanità. Tra queste attività, quella della giustizia e quella della medicina, sono certamente al centro del dibattito istituzionale europeo. Nel contesto medico, che approfondiremo come caso di studio nel successivo capitolo, caratterizzato da una relazione di fiducia tra medico e paziente, così come nel contesto della giustizia, caratterizzato, a sua volta, dalla richiesta di valutazione di delicati interessi, soprattutto nel settore penalistico, il problema dell'esplicabilità dell'algoritmo si pone sia quando è usato come metodo per filtrare i dati sia soprattutto quando lo stesso algoritmo diviene un modo per esternalizzare il processo decisionale dall'uomo alla macchina o a un *software bot*.

Voci di corridoio nel dibattito istituzionale sostengono che il principio di esplic-

⁶³ Ad esempio, M.E. KAMINSKI, *Understanding Transparency in Algorithmic Accountability*, in W. BARFIELD (eds.) *The Cambridge Handbook of The Law of Algorithms*, Cambridge, Cambridge University Press, 2021, pp. 121-138.

⁶⁴ Per una ricognizione analitica dei documenti europei in materia cfr. F.H. LLANO-ALONSO, *L'etica dell'intelligenza artificiale nel quadro giuridico dell'Unione Europea*, cit.

⁶⁵ Cfr. ad esempio, PANEL FOR THE FUTURE OF SCIENCE AND TECHNOLOGY (STOA), *A Governance Framework for Algorithmic Accountability and Transparency*, 2019, accessibile al link [https://www.europarl.europa.eu/thinktank/en/document/EPRS_STU\(2019\)624262](https://www.europarl.europa.eu/thinktank/en/document/EPRS_STU(2019)624262).

cabilità e la correlata comprensibilità verranno di fatto abbandonati quando l'uso degli algoritmi andrà a regime per diverse attività e pertanto guadagnerà con la diffusione il crisma dell'affidabilità. In siffatto scenario, difficilmente il professionista – giudice o medico – si discosterà dall'output indicato dal sistema.

In pratica, così come per viaggiare è diventato d'uso utilizzare il navigatore e non la tradizionale mappa cartacea, allo stesso modo ci si adatterà a delegare una parte o, addirittura, tutta la fase decisionale a sistemi di IA, senza più porsi il problema se quello che propongono sia la soluzione migliore nel caso di specie. Come è stato sottolineato «ogniqualevolta un automatismo decisionale viene inserito in un procedimento deliberativo, l'automatismo tende a “catturare” la decisione, o quantomeno a rendere estremamente difficile prescindere»⁶⁶.

Questa deriva andrebbe evitata, soprattutto nei settori poc'anzi richiamati, per almeno due ragioni. La prima riguarda il rispetto della professionalità degli esperti, che potrebbero voler, anche senza dovere, conoscere i contenuti di ciò a cui affidano parte della loro decisione. La padronanza e la conoscenza dell'IA contribuirebbe a mantenere alto il profilo dell'autorealizzazione personale del professionista attraverso il proprio lavoro, che esige spirito critico e attenzione al miglior interesse di chi si affida all'esperto per una decisione sulla sua esistenza – medico o giudice. Inoltre, evitare la deriva del cieco affidamento consentirebbe di non deresponsabilizzare il decisore umano, che potrebbe sentirsi sollevato «dal peso e dal rischio di dover motivare e rispondere di quelle motivazioni»⁶⁷. Contrariamente al pensiero diffuso in alcune prospettive a favore di una sostituzione dell'attività umana con la tecnologia, la volontà di mantenere il controllo e l'assunzione di responsabilità potrebbe essere un buon argomento per salvaguardare alcune professioni specialistiche – il radiologo, l'ortopedico, ad esempio – che rischiano altrimenti di sparire. Chi scrive ritiene che non si siano ancora completamente chiariti né gli svantaggi della perdita di queste e di altre professionalità⁶⁸ né la capacità del settore dell'IA di riuscire a creare *decent jobs*⁶⁹, in alternativa a quelli che sono destinati a essere cancellati dall'uso dell'IA.

La seconda ragione si rivolge ai destinatari delle decisioni – mediche o giudiziali – assunte da sistemi di IA. Non poter decidere di conoscere i contenuti rilevanti del processo, che ha portato un sistema artificiale a decidere per le loro vite, li mette in una posizione di subordinazione totale e non tanto rispetto al

⁶⁶ 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, Giappichelli, 2020, pp. 1-42, ivi, p. 39.

⁶⁷ *Ibidem*, p. 40.

⁶⁸ Uno studio dettagliato in tema di lavoro richiesto dall'AIDA Committee, Parlamento europeo, intitola *Improving Working Conditions using Artificial Intelligence*, 2021, accessibile a questo link [https://www.europarl.europa.eu/RegData/etudes/STUD/2021/662911/IPOL_STU\(2021\)662911_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2021/662911/IPOL_STU(2021)662911_EN.pdf).

⁶⁹ *Ibidem*.

decisore umano quanto a quello macchinico. Questa conseguenza non solo minerà la fiducia nei sistemi non controllabili dall'uomo, ma condurrà a derive disumanizzanti perché la vita di chi è colpito da queste decisioni *senza contraddittorio* è costantemente a rischio di discriminazione⁷⁰.

Si ritiene, inoltre, che sia importante non abbandonare il principio di esplicabilità, perché, in coerenza con il principio di precauzione, essendo agli inizi di un percorso di trasformazione di una vasta gamma di attività umane, non siamo ancora in grado di prevedere obiettivamente i potenziali rischi di una cessione completa della nostra decisionalità alle macchine.

Quale ruolo centrale questo principio può giocare nel contesto medico sarà oggetto di approfondita analisi nel successivo capitolo.

Per quello che riguarda il contesto della giustizia, ci limitiamo ad osservare che poter comprendere l'esatto scopo di un algoritmo rappresenta un fattore centrale per valutare l'estensione della violazione del diritto da parte di quel dato sistema⁷¹. Nel contesto di *Civil Law*, i giudici devono poter comprendere quello che devono valutare, per quanto complesso questo sia. Le perizie degli esperti in una data materia scientifica sono l'utile strumento di chiarificazione al giudice della portata di una data prova scientifica. Ma il giudice resta il *gatekeeper* della valutazione finale. Ciò deve persistere anche con riferimento ai sistemi di IA. Dovranno essere gli informatici a spiegare il funzionamento di complessi algoritmi, ma sarà il giudice a stabilire quanto debba pesare il problema tecnico nella decisione finale quando in gioco ci sono valori come l'integrità personale, la libertà, l'eguaglianza di trattamento e la dignità umana. Una proposta che dovrebbe venire seriamente valutata e modulata per le esigenze specifiche del settore giustizia è quella di costituire corti specializzate in dispute dove sono coinvolti algoritmi. «Perhaps in response to the complexity of algorithms, a court which focuses on litigation involving algorithms will need to be developed. This is not a novel idea as in the United States there is a Tax Court, a Court of Federal Claims [...] to name just a few courts with specialized subject matter»⁷².

Si tratta di una proposta da valutare, se ciò consente di proteggere in maniera più incisiva i diritti potenzialmente violabili a mezzo di IA.

Va inoltre considerato che per il settore della giustizia, la richiesta di trasparenza e di controllo da parte dell'utente è stata formulata espressamente anche dalla *European Commission for the Efficiency of Justice* nel 2018 nell'*European*

⁷⁰ A. SIMONCINI, *Amministrazione digitale algoritmica. Il quadro costituzionale*, cit. p. 37, «se l'intervento umano serve a rendere conoscibile la motivazione e trasparente l'iter logico della decisione automatizzata, entrambi questi elementi sono decisivi per determinare in concreto l'esistenza o meno di una ragione discriminatoria o di una violazione del principio di eguaglianza».

⁷¹ Insistono su questo punto W. BARFIELD, J. BARFIELD, *An Introduction to Law and Algorithms*, in W. BARFIELD (eds.), *The Cambridge Handbook of The Law of Algorithms*, Cambridge, Cambridge University Press, 2021, pp. 3-15.

⁷² *Ibidem*, p. 7.

*ethical Charter on the use of Artificial Intelligence in judicial systems and their environment*⁷³. Il documento prevede cinque principi guida⁷⁴, di cui il quarto e il quinto richiedono di rendere i metodi di processamento dei dati accessibili e comprensibili e autorizzare *audit* esterni (principio di trasparenza, imparzialità e *fairness*), nonché vietare approcci prescrittivi e assicurare che gli utenti siano attori informati, in grado di controllare le scelte che fanno (principio *under user control*). Si esplicita, inoltre, che «[a] judicial system in keeping with its time would be one that is capable of establishing, administering and guaranteeing genuine cyberethics for both the public and private sectors, and insisting on total transparency and fairness in the functioning of algorithms, which may contribute one day to judicial decision-making»⁷⁵.

⁷³ Il documento è accessibile al link <https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c>.

⁷⁴ I cinque principi sono: 1) Principle of respect for fundamental rights: ensure that the design and implementation of artificial intelligence tools and services are compatible with fundamental rights; 2) Principle of non-discrimination: Specifically prevent the development or intensification of any discrimination between individuals or groups of individuals; 3) Principle of quality and security: With regard to the processing of judicial decisions and data, use certified sources and intangible data with models conceived in a multi-disciplinary manner, in a secure technological environment; 4) Principle of transparency, impartiality and fairness: Make data processing methods accessible and understandable, authorise external audits; 5) Principle “under user control”: Preclude a prescriptive approach and ensure that users are informed actors and in control of their choices.

⁷⁵ *Ibidem*, p. 59.

CAPITOLO IV

CASO DI STUDIO: IA E *BIG DATA* IN MEDICINA

SOMMARIO: 1. Introduzione. – 2. I dati e loro impiego nella pratica clinica e nella ricerca scientifica. – 3. IA e relazione di cura: terzo *super partes*? – 3.1. La relazione di cura tra diritti e doveri. – 3.2. La relazione di cura mediata dall'IA. – 4. Neurotecnologie e potenziamento umano: le nuove frontiere della medicina. – 4.1. Il biopotenziamento in breve. – 4.2. Profili etico-giuridici del potenziamento neurotecnologico. – 4.3. Neurotecnologie e ideologia antropomorfizzante: alcune riflessioni conclusive.

1. *Introduzione*

In questo capitolo si porterà l'attenzione sull'uso dell'IA in medicina, come caso di studio di alcune questioni trattate nei capitoli precedenti. La dialettica degli opposti e l'ideologia antropomorfizzante sono sottese all'impiego dell'IA anche in questo specifico contesto e, pertanto, vale la pena indagarne gli effetti.

Si avrà modo di analizzare criticamente le applicazioni dell'IA sia nella pratica clinica sia nella ricerca clinica. In tali ambiti, svariate sono sia le possibilità di impiego di queste tecnologie sia le possibili problematiche etiche e le conseguenti soluzioni giuridiche¹.

Mi riferisco, innanzitutto, alle questioni sollevate dall'uso dei dati e dalle concezioni sottese a questo uso. Nei capitoli precedenti si è dato conto della diffusione di una ideologia *essenzialistica* e deterministica che accompagna la nar-

¹ Si trovano riferimenti ai molteplici impieghi dell'IA in medicina in S. SALARDI, M. SAPORITI, *Dai principi alle regole. Appunti per una riflessione sul rapporto tra bioetica e intelligenza artificiale*, in S. SALARDI, M. SAPORITI, M. VETIS ZAGANELLI (a cura di), *Diritti umani e tecnologie morali. Una prospettiva comparata tra Italia e Brasile-Direitos Humanos e tecnologias morais. Uma perspectiva comparada entre Itália e Brasil*, Torino, Giappichelli, 2022, pp. 101-117. Si possono distinguere applicazioni dell'IA a fini diagnostici, terapeutici, ad esempio per le diagnosi radiologiche in oncologia, in campo dermatologico, o per diagnosi di retinopatia diabetica, così come nel contesto oftalmologico e immunologico. Vi sono poi applicazioni che riguardano gli aspetti operativi e amministrativi, nonché quelle usate per la sorveglianza della salute pubblica.

razione dell'IA e dei dati, di cui quest'ultima si alimenta. Mutuando le parole di Stefano Rodotà², potremmo dire che si è venuta consolidando un'adesione *mistica* all'ideologia dei dati, radicata in un'*infinite naivetè*³ circa la portata dei risultati ottenuti dal trattamento dei dati. Sebbene molte siano le voci critiche nei confronti di questa visione riguardante i dati, è radicata l'idea che i dati rappresentino l'unica fonte su cui fare assoluto affidamento, in quanto rappresenterebbero in maniera completa il reale. Questa concezione ha parecchie ricadute problematiche quando i dati sono usati in medicina, in quanto il cieco affidamento a essi rischia di depotenziare il ruolo del medico e di rimettere in discussione il modello informativo attualmente vigente. Inoltre, le conseguenze di questa *datadeology* si fa strada nella relazione di cura sollevando, ad esempio, problemi di *dipendenza* della decisione medica dalle tecnologie intelligenti, limitando sempre più la possibilità per il professionista di utilizzare le proprie competenze e conoscenze come criterio di verifica di quello che l'analisi algoritmica gli mette comodamente a disposizione a fini diagnostici, prognostici, terapeutici ecc. Anche in questo ambito uno spinoso problema è quello di riuscire a mantenere una distanza critica rispetto all'adozione di tecniche e tecnologie, che sembrano dover essere necessariamente accettate, senza porsi troppi quesiti, in quanto *fanno il bene dell'umanità salvando vite*.

In aggiunta a quanto sinora evidenziato, va sottolineato che anche la medicina risente della narrazione antropomorfizzante dell'IA, di cui abbiamo dato conto in precedenza e che qui esemplificheremo.

In breve, l'ambito della medicina rappresenta un osservatorio privilegiato dei cambiamenti che la digitalizzazione delle attività umane può apportare alla nostra *relazionalità*. Quello della relazione di cura è, infatti, uno dei contesti relazionali per eccellenza, di cui, sin dall'epoca di Ippocrate, si è cercato di definire i confini, i contenuti e il ruolo degli attori coinvolti in merito alla spettanza delle decisioni.

Questa relazionalità si esprimeva in origine nella relazione tra medico e paziente. Solo con la riflessione bioetica, originata nel secolo scorso da sviluppi in campo scientifico e tecnologico, la relazione è arrivata a includere anche soggetti appartenenti all'*équipe*⁴. A ognuno di questi soggetti sono stati riconosciuti ruoli, competenze, poteri, doveri e responsabilità da un punto di vista giuridico, oltre che etico e deontologico. In questo complesso scenario, caratterizzato dall'intrecciarsi di posizioni etiche e giuridiche molto articolate, si colloca oggi l'im-

²Stefano Rodotà usava il termine *mistica del DNA*, cfr. S. RODOTÀ, *Se la società impone la schedatura genetica*, 2005 reperibile al sito www.privacy.it, ultimo accesso settembre 2022.

³P. KIRCHSCHLAEGER, *Digital Transformation and Ethics*, Baden-Baden, Nomos, 2021, p. 298.

⁴Il medico mantiene la responsabilità informativa, la competenza di prospettare scenari e alternative di intervento, mentre al paziente spetta la decisione sul se sottoporsi o meno ai trattamenti proposti. Infine, i componenti dell'*équipe*, in particolare gli infermieri hanno un dovere di assistenza e una correlata autonomia relativa ai compiti deontologicamente loro assegnati.

piego dell'IA. La presenza di questo terzo ente modificherà, nel tempo, la relazionalità, che caratterizza questo rapporto. Ciò può avvenire o attraverso un contributo dell'IA virtuoso e migliorativo dell'attività del medico o, al contrario, inducendo una spersonalizzazione della relazione di cura. Il *dual-use*, che è una peculiarità intrinseca di queste tecnologie, richiede una attenta valutazione degli effetti che produrrà il loro impiego, sul lungo termine, nel contesto della medicina. Bisogna, quindi, ponderare bene la scelta degli scopi dell'IA in medicina.

La trasformazione digitale di questa attività porta con sé, infatti, rischi di violazione dei diritti fondamentali come quello all'identità personale, alla non discriminazione, al libero sviluppo della propria personalità, alla privacy e alla dignità: rischi che potrebbero compromettere i vantaggi di una medicina personalizzata. La personalizzazione della medicina, volta a rendere l'esercizio del diritto alla salute più effettivo, presenta sempre un'altra faccia: quella del controllo sulle persone. La personalizzazione della medicina può essere costruita su due modelli. Il primo è quello in cui il medico è presente ed è supportato dalle tecnologie intelligenti per fornire una risposta sempre più individualizzata e corretta. Il secondo è quello in cui il medico è assente, o meglio diviene un mero esecutore delle decisioni dell'IA. Questo secondo modello, che è sostenuto, più o meno esplicitamente da molti detentori del potere tecnologico ispirati dalla filosofia transumanista, rappresenta una sfida etica per la relazionalità al centro di questa attività. Trattasi di un rapporto relazionale che fa della mutua fiducia un elemento centrale su cui costruire una virtuosa collaborazione tra pazienti e professionisti della salute, nel miglior interesse del paziente.

Le riflessioni che seguono sono volte a mettere in luce le criticità qui brevemente richiamate.

2. I dati e loro impiego nella pratica clinica e nella ricerca scientifica

L'ambito medico rappresenta il contesto più produttivo di dati personali altamente sensibili: dai dati sulla salute ai dati biometrici e genetici⁵, fino ai dati

⁵ All'articolo 4, il Regolamento europeo sulla protezione dei dati (GDPR) 679/2016 definisce:

Dati relativi alla salute: i dati personali attinenti alla salute fisica o mentale di una persona fisica, compresa la prestazione di servizi di assistenza sanitaria, che rivelano informazioni relative al suo stato di salute.

Dati biometrici: i dati personali ottenuti da un trattamento tecnico specifico relativi, alle caratteristiche fisiche, fisiologiche o comportamentali di una persona fisica che ne consentono o confermano l'identificazione univoca, quali l'immagine facciale o i dati dattiloscopici.

Dati genetici: i dati personali relativi, alle caratteristiche genetiche ereditarie o acquisite di una persona fisica che forniscono informazioni univoche sulla fisiologia o sulla salute di detta persona fisica, e che risultano in particolare dall'analisi di un campione biologico della persona fisica in questione.

derivanti dagli studi e dalle applicazioni neuroscientifiche (mappatura del cervello). Questi dati sono fondamentali per migliorare la pratica clinica e la ricerca scientifica in campo biomedico⁶. Oggigiorno, la quantità di dati gestiti è talmente grande, che è stato coniato il termine *Big Data* per dare l'idea del contesto entro cui ci si muove. In generale, va ricordato che, in virtù delle loro caratteristiche, i *Big Data* sono impossibili da gestire seguendo i tradizionali principi di trattamento dei dati e richiedono l'individuazione di procedure adeguate a salvaguardare la tutela dei diritti fondamentali e delle informazioni personali. Sebbene vi siano diverse definizioni di *Big Data*, che variano a seconda della disciplina, vi è una generale concordanza sul fatto che sono caratterizzati dalle c.d. "4V": volume, velocità, varietà, veracità⁷. Abbiamo già analizzato nel capitolo secondo queste caratteristiche e anche le loro problematicità etiche, rimandiamo, pertanto, alle osservazioni ivi elaborate. In questa sede, ci concentreremo sul loro impiego in medicina.

I *Big Data* sono masse di dati eterogenee, che derivano sia direttamente dal settore medico (dalle analisi genetiche, dalle immagini radiologiche, dalle cartelle cliniche ecc.) sia indirettamente da altri settori. In questo caso si tratta di dati convertibili in dati relativi alla salute (ad esempio, dati sullo stile di vita, sulle condizioni ambientali, socioeconomiche, comportamentali e sociali). L'Organizzazione Mondiale della Sanità considera i *Big Data* biomedici costitutivi di un ecosistema di dati sanitari (*health data ecosystem*)⁸. Sono dati soggetti a trattamenti automatizzati di carattere algoritmico, che vengono impiegati per addestrare i software di IA, che saranno in grado, a fine addestramento, di fornire soluzioni diagnostiche e prognostiche a supporto delle decisioni mediche o di creare databases a scopo di ricerca scientifica.

L'uso di questi dati solleva problemi etico-giuridici che richiedono un'articolata analisi a seconda del loro impiego nell'ambito della pratica clinica o della ricerca scientifica.

Nella pratica clinica il potenziale offerto dalla disponibilità di questa quantità enorme di dati spiega i suoi effetti nell'ambito della c.d. *data-driven precision*

⁶ Per un approfondimento sul tema della protezione dei dati nell'ambito della ricerca scientifica cfr. F. DI TANO, *Protezione dei dati personali e ricerca scientifica: un rapporto controverso ma necessario*, in *Biolaw Journal-Rivista di Biodiritto*, 2022, 1, pp. 71-98.

⁷ Vi è chi considera una quinta V, corrispondente al concetto di *valore*. Questo è inteso come «la rilevanza e il significato dei dati nel contesto attuale, con riferimento all'utente o anche alle compagnie che li raccolgono», L. PALAZZANI, *Tecnologie dell'Informazione e Intelligenza Artificiale. Sfide Etiche al Diritto*, Roma, Studium edizioni, 2020, p. 27. Cfr. UNESCO International Bioethics Committee, *Report of the IBC on Big Data and Health*, 2017, che richiama le 5V.

⁸ WORLD HEALTH ORGANIZATION, *Ethics and Governance of Artificial Intelligence for Health*, 2021, p. 35, accessibile al seguente link <https://www.who.int/publications/i/item/9789240029200>, ultimo accesso agosto 2022. Le fonti degli *health data* sono moltissime e sono ben rappresentate nel documento menzionato in un grafico esaustivo.

medicine, intesa come «la possibilità [...] di costruire sulla base dei dati raccolti (sulla salute ma non solo), la predizione e la simulazione virtuale di strategie di intervento per singoli pazienti in base alle specifiche caratteristiche nel contesto della medicina personalizzata, stratificata e di precisione»⁹.

Nonostante le straordinarie opportunità offerte dai *Big Data*, non si può nascondere che il loro impiego sollevi diversi problemi etici. Nel secondo capitolo ci siamo occupati in generale di alcune di queste criticità. Di seguito vogliamo soffermarci su quelle relative all'uso dei *Big Data* quando sono le decisioni mediche a basarsi sugli output generati dall'IA. Come noto per elaborare, gestire, classificare e trattare i dati di queste dimensioni l'avvento dell'IA algoritmica è stato decisivo. Partendo dal presupposto che le tecnologie di IA rimangano un supporto alla decisione medica e non la sostituiscano¹⁰, possiamo individuare i principali nodi critici derivanti dai *Big Data* che le alimentano. Se analizzati sul piano tecnico, molti di questi problemi si configurano come criticità, che nel tempo verranno, in buona parte, risolte con l'avanzamento tecnologico. Tuttavia, dietro questi problemi tecnici si celano precisi assunti etici circa la spettanza delle decisioni. Se è vero che il medico può contare oggi su un ulteriore strumento per affinare le sue diagnosi, è altresì vero che i problemi derivanti dalle 4V dei *Big Data* non sono ancora risolti e possono incidere, anche non marginalmente, sull'accuratezza della diagnosi e sulla qualità della cura. Paradossalmente, in contrasto con le aspettative di migliorare la velocità, l'accuratezza e ridurre la componente soggettiva del processo decisionale, le decisioni supportate da IA potrebbero indurre errori diagnostici e terapeutici, non dovuti a imperizia, imprudenza o negligenza del medico.

Garante, come deve essere, di una corretta e veritiera informazione al paziente, il medico deve, infatti, poter contare su *outputs* che siano privi di *bias*. Per ottenere questo risultato, occorre che i dati siano "puliti" e di alta qualità (*high-quality data*). Se questo non fosse il caso, anche nell'ipotesi da cui siamo partiti, ossia che l'IA assista l'attività decisionale medica e non la sostituisca, di fatto lo farà implicitamente, perché il medico dovrà fondare le sue decisioni tenendo conto, dove non può farne a meno, di risultati su cui non ha alcun controllo. Ciò inficia alla base la reale spettanza delle decisioni e il dovere di corretta prospettiva dei quadri diagnostici e prognostici. I problemi della composizione dei

⁹L. PALAZZANI, *Tecnologie dell'Informazione e Intelligenza Artificiale. Sfide Etiche al Diritto*, cit. p. 27.

¹⁰Si tratta della *human-centric vision* auspicata dall'Unione europea con i suoi provvedimenti normativi e condivisa dal Consiglio d'Europa.

Esempi di questo tipo di uso si trovano nell'ambito radiologico, dove l'IA algoritmica applicata alle mammografie assiste i radiologi nell'individuazione del cancro al seno. O ancora in ambito psichiatrico, avatar virtuali avanzati tengono conversazioni che hanno ricadute per la diagnosi e il trattamento psichiatrico. Cfr. M.J. RIGBY, *Ethical Dimensions of Using Artificial Intelligence in Health Care*, in *AMA Journal of Ethics*, 2019, 21, 2, pp. E 121-124.

dati inseriti nei *datasets* sono noti da qualche anno. Si tratta, innanzitutto, di garantire la completezza e la qualità dei dati con cui si addestra l'IA, in generale, ma a maggior ragione, quando questi dati riguardano, e al contempo derivano da, paesi a basso e medio reddito, caratterizzati da una sottorappresentazione endemica di certe categorie di soggetti. Inoltre, i dati con cui si addestra l'IA contengono ancora *bias* sistemici, dovuti alla sottorappresentazione di interi gruppi di persone. Ai problemi, direttamente connessi ai dati, si aggiungono quelli derivanti dai modelli predittivi algoritmici, minati dall'opacità e dalla limitata esplicabilità (*explainability*) circa il loro funzionamento.

Questi problemi tecnici celano problemi etico-giuridici, in quanto sono alla base di possibili violazioni di una serie di principi e di diritti fondamentali: dall'identità personale, al diritto alla non discriminazione e all'eguaglianza di trattamento, al diritto alla privacy, al diritto all'autonomia informativa, incluso il diritto di non sapere¹¹. In generale, i problemi evidenziati indeboliscono la visione incentrata sulle persone, perché rendono, di fatto, impossibile il *meaningful human control*¹². Questo concetto è stato elaborato per affrontare la spinosa questione dell'attribuzione della responsabilità, invero se il controllo umano (medico) è impossibile, diviene complesso attribuire la responsabilità per diagnosi errate.

Un ulteriore profilo di criticità derivante dalla trasformazione della medicina in *medicina guidata dai dati* è la difficoltà di restare immune dal *dataismo*, che abbiamo definito nel secondo capitolo. Anche il contesto clinico non è infatti esente dal pericolo di ridurre gli esseri umani a categorie di rischio e di salute, a detrimento della diversità individuale. Con l'attenzione incentrata sui dati, si sposta lo sguardo della medicina dal soggetto in carne ed ossa ai suoi dati, ai numeri e ai modelli, con il rischio di trasformare la medicina personalizzata in

¹¹ Un diritto che si è affermato con forza in occasione del progresso scientifico in ambito genetico, con importanti previsioni normative, come la Convenzione sui diritti dell'Uomo e la Biomedicina del Consiglio d'Europa del 1998. Duttge Gunnar vede il rafforzamento del diritto a non sapere come un importante passo per garantire all'individuo la sua sfera privata. «Wenn der Patient bzw. Proband in dieser Moderne mehr und mehr nur noch als „Knotenpunkt der Datenströme“ erscheint, bietet ein verstärkter Ausbau des Rechts auf Nichtwissen eine Chance, sich dennoch hier und da noch eine Nische des Privaten zu erhalten», cfr. G. DUTTGE, *Das Recht auf Nichtwissen in einer informationell vernetzten Gesundheitsversorgung*, in *Medizinrecht*, 2016, 34, pp. 664-669, ivi, ivi, p. 669.

¹² Cfr. sul concetto F. SANTONI DE SIO, J. VAN DE HOVEN, *Meaningful Human Control over Autonomous Systems: A Philosophical Account*, in *Frontiers in Robotic and AI*, 2018, 5, pp. 1-14. Anche diversi comitati etici nazionali si sono espressi a favore del mantenimento del "controllo umano significativo", cfr. tra gli altri il Comitato Nazionale per la Bioetica (CNB) e il Comitato Nazionale per la Biosicurezza, le Biotecnologie e le Scienze della Vita (CNBBSV) che nel parere del Gruppo misto del 2020 intitolato *Intelligenza Artificiale e Medicina: Aspetti Etici* tengono fermo questo punto. In tale direzione, anche il Nuffield Council on Bioethics nel parere dal titolo *Artificial intelligence (AI) in healthcare and research* del 2020; The Swedish National Council of Medical Ethics, *In brief-Artificial Intelligence in healthcare*, 2020; Deutscher Ethikrat, *Big Data und Gesundheitsdatensouveränität als informationelle Freiheitsgestaltung*, 2017.

*dehumanized o depersonalized medicine*¹³. Si rischia, quindi, anche in questo delicato settore, una tirannia dei dati che può tradursi in altrettanti comportamenti oppressivi di controllo e monitoraggio delle persone, oltre i limiti dettati dalle esigenze di salute individuale e pubblica.

Se, dalla pratica clinica, ci spostiamo ad analizzare il contesto della ricerca scientifica, ci avvediamo che molti dei problemi delineati in merito alla qualità, all'accuratezza, alla validità e alla disponibilità di dati clinici completi si ripropongono anche in questo contesto. In questo ambito, diviene centrale poter accedere alle raccolte di dati esistenti per costruire ulteriori *datasets*. Rispetto a questi ultimi si pone, innanzitutto, il problema tecnico di accesso, in particolare quando i *datasets* si trovano in paesi extra-UE o viceversa, quando dall'Unione europea devono essere trasmessi i propri *datasets* a realtà con normative molto lontane dai contenuti stringenti del Regolamento europeo sulla protezione dei dati (GDPR, Reg. 679/2016). Correlato alla questione tecnica dell'accesso, che pone rischi di furto o di diffusione accidentale di dati¹⁴, vi è, infatti, la questione etico-giuridica di garanzia di un uso trasparente dei dati per limitare i rischi di discriminazione, in particolare di categorie vulnerabili, come i minori. Questi ultimi potrebbero essere oggetto di stigmatizzazione da adulti, a causa dei dati accumulati su di loro quando erano minorenni. Questi rischi sono diventati più pressanti da quando la pandemia Covid-19 ha accelerato la costituzione di infrastrutture per raccogliere e conservare i dati sulla salute per usi molto diversi fra loro¹⁵. Questi *datasets* usati per addestrare l'IA sono tuttora fonte di problematicità, in quanto non c'è certezza che siano scevri di *biases*.

In breve, l'uso di grandi quantità di dati nella pratica clinica e nella ricerca biomedica apre straordinarie opportunità, fino a qualche decennio fa inimmaginabili. Tuttavia, a queste opportunità si affiancano rilevanti quesiti etici, che non hanno ancora trovato una corretta e compiuta definizione giuridica. La Proposta di Regolamento Europeo sull'IA si pone come primo strumento vincolante per gli usi dell'IA nella società europea. Tale Proposta non contempla esplicitamente il campo medico. Per comprendere le ricadute di tale normativa sulla medicina occorrerà attendere il Regolamento che verrà realizzato. Se la Proposta si tradurrà in un Regolamento realmente incentrato sulle persone, potrà certamente svolgere un importante ruolo di tutela dei soggetti anche in campo medico, limitando il rischio di una medicina deumanizzata.

¹³ Cfr. P. KIRCHSCHLAEGER, *Digital Transformation and Ethics*, cit. p. 216.

¹⁴ Già da un decennio si punta l'attenzione sui rischi di malfunzionamento dei sistemi di IA nel contesto statunitense, in particolare quelli che determinano fughe di dati esponendo i pazienti a gravi pericoli, cfr. D.M. DUDZINSKI, P.C. HÉBERT, M.B. FOGLIA, T.H. GALLAGHER, *The disclosure dilemma. Large-scale adverse events*, in *New England Journal of Medicine*, 210, 363, 10, pp. 978-986.

¹⁵ Nelle linee guida dell'OMS, *supra* nota 8, si ricorda il sistema QR cinese elaborato da un sistema di pagamento digitale di Alipay. Si è istituito un "Alipay Health Code" che consente di dedurre dai dati inseriti se una persona è a rischio di contagio.

I rischi di una medicina deumanizzata si palesano con particolare riguardo alla relazione di cura. Analizzeremo questa questione nel prossimo paragrafo.

3. IA e relazione di cura: terzo super partes?

Nello *Strategic Action Plan on Human Rights and Technologies in Biomedicine (2020-2025)* del novembre 2019, il *Committee on Bioethics (DH-BIO)* del Consiglio d'Europa¹⁶ pone tra i suoi obiettivi quello dell'elaborazione di un rapporto sulle applicazioni dell'IA nell'ambito sanitario, con particolare attenzione rivolta alla relazione di cura¹⁷. A giugno 2022, lo *Steering Committee for Human Rights in the field of Biomedicine and Health (CDBIO)* ha pubblicato il rapporto intitolato *The Impact of Artificial Intelligence on the Doctor-Patient Relationship*¹⁸. Il Comitato ha analizzato, in maniera approfondita, le ricadute dell'IA nell'ambito della relazione di cura, evidenziando le maggiori criticità. Includendo le riflessioni entro i confini della cornice dei principi e delle norme previste dalla Convenzione sui Diritti dell'Uomo e la Biomedicina (Convenzione di Oviedo) in materia di relazione di cura, il Comitato si è concentrato su considerazioni relative all'IA applicata alla ricerca biomedica e ai pazienti¹⁹. Sebbene nel rapporto si sottolinei che l'impiego dell'IA nel contesto clinico sia ancora agli albori, si propone un'analisi della potenziale trasformazione della re-

¹⁶Dal 1° gennaio 2022 è divenuto lo Steering Committee for Human Rights in the field of Biomedicine and Health (CDBIO), consultare il seguente link per maggiori dettagli [https://www.coe.int/en/web/bioethics/-/the-committee-on-bioethics-dh-bio-becomes-the-steering-committee-for-human-rights-in-the-fields-of-biomedicine-and-health-cdbio-#:~:text=Back-,The%20Committee%20on%20Bioethics%20\(DH%20DBIO\)%20becomes%20the%20Steering,of%20Biomedicine%20and%20Health%20\(CDBIO\)&text=On%202024%20November%202021%2C%20the,quadrennium%20\(2022%20%E2%80%93%202025\)](https://www.coe.int/en/web/bioethics/-/the-committee-on-bioethics-dh-bio-becomes-the-steering-committee-for-human-rights-in-the-fields-of-biomedicine-and-health-cdbio-#:~:text=Back-,The%20Committee%20on%20Bioethics%20(DH%20DBIO)%20becomes%20the%20Steering,of%20Biomedicine%20and%20Health%20(CDBIO)&text=On%202024%20November%202021%2C%20the,quadrennium%20(2022%20%E2%80%93%202025)).

¹⁷Così si esprime il Comitato «Artificial Intelligence (AI) has the potential to improve diagnostic and therapeutic outcomes for patients. Although deep learning algorithms in a variety of tasks in radiology and in medicine generally have demonstrated significant promise, it is likely to be several years before AI is mainstreamed into the healthcare domain. The predictive capability of AI raises concerns about privacy and discrimination. Moreover, as AI evolves, it will create new complexities for the doctor-patient relationship. In the light of these challenges, the Committee on Bioethics intends to prepare a report highlighting the role of healthcare professionals in respecting the autonomy, and right to information, of the patient, and in maintaining transparency and patient trust as critical components of the therapeutic relationship», cfr. COMMITTEE ON BIOETHICS (DH-BIO), *Strategic Action Plan on Human Rights and Technologies in Biomedicine (2020-2025)*, 2019.

¹⁸STEERING COMMITTEE FOR HUMAN RIGHTS IN THE FIELD OF BIOMEDICINE AND HEALTH (CDBIO), *The Impact of Artificial Intelligence on the Doctor-Patient Relationship*, 2021, accessibile a questo link <https://www.coe.int/en/web/bioethics/developing-a-report-on-the-application-of-ai-in-healthcare-in-particular-regarding-its-impact-on-the-doctor-patient-relationship>.

¹⁹*Ibidem*, p. 29.

lazione di cura mediata dall'IA (*the AI-mediated doctor-patient relationship*). Basandoci su alcuni elementi contenuti nel rapporto, ma con l'intento di appor- tare qualche ulteriore riflessione, ci concentreremo, principalmente anche se non esclusivamente, sull'ipotesi in cui i sistemi di IA svolgano un ruolo nel pro- cesso decisionale del medico, ad esempio, fornendo raccomandazioni per la dia- gnosi.

Prima di procedere in tal senso, appare opportuno richiamare gli aspetti fon- danti della relazione di cura così come si è venuta configurando dalla seconda metà del secolo scorso ad oggi, con il prezioso contributo della riflessione bioe- tica. Questa operazione è essenziale per comprendere su quali aspetti di questa relazione può incidere l'IA.

3.1. *La relazione di cura tra diritti e doveri*

Sebbene oggigiorno siamo abituati a parlare di diritti dei pazienti, della loro autonomia decisionale relativamente alla propria salute e di diritto di sapere (con il correlato diritto di non sapere), questa dimensione non è una conquista che risale a molto tempo fa²⁰. I diritti richiamati sono frutto delle battaglie per i diritti, iniziate nella seconda metà del secolo scorso nel contesto occidentale. Queste battaglie hanno visto coinvolte diverse categorie di soggetti storicamen- te discriminate sul piano dell'autonomia decisionale – donne, minori, persone appartenenti a particolari etnie – e che, a un certo punto della storia, hanno ri- vendicato la possibilità di decidere per sé o, quanto meno, di partecipare ai processi che li riguardano attraverso il coinvolgimento sul piano informativo²¹. Entro i confini di queste battaglie si è inserita quella dei pazienti, iniziata nel contesto statunitense tra la fine degli anni Sessanta e l'inizio degli anni Settanta del secolo scorso, si è poi estesa agli altri paesi occidentali. È una concomitanza di condizioni ad avere consentito ai pazienti di vincere questa battaglia. Si tra- ta di condizioni culturali (affermazione del principio di autonomia) ed econo- miche (diffusione del benessere economico a larghi strati della popolazione) che, unitamente all'inedito progresso scientifico e tecnologico, hanno giocato un ruolo determinante nel successo delle rivendicazioni dei pazienti. In poche parole, si fa strada l'idea che, nell'era della tecnologia, che ha dato alla medici- na un inedito potere di controllare processi vitali e biologici, un tempo lasciati

²⁰ Come è stato osservato «il rapporto medico-paziente si è strutturato come espressione delle culture e delle società che lo hanno prodotto [...] esso si è articolato in forme di comunicazione diverse, alternativamente secondo rapporti di simmetria e di asimmetria e secondo forme variabili di autorità e collaborazione», cfr. L.R. ANGELETTI, V. GAZZANIGA, *Storia, filosofia ed etica genera- le della medicina*, quarta ed., Milano, Elsevier, 2012, p. 1.

²¹ È questo certamente il caso dei minori, che in maniera sempre più definita sul piano giuridi- co, hanno ottenuto il diritto di essere parte del processo informativo nell'ambito delle vicende che li riguardano. Cfr. per tutti in materia sanitaria l'articolo 3 della legge n. 219/2017.

in balia del caso²², la spettanza delle decisioni, da tempo risalente di indiscussa competenza del medico (modello paternalistico della relazione di cura), non possa più essere data per scontata.

A fronte dell'immenso potere di cui il medico dispone, il mandato della sua attività non può più essere inteso solo sul piano tecnico-scientifico. Nel nuovo scenario tecnologico, tale prospettiva appare insufficiente a giustificare il suo esclusivo potere decisionale. Dietro tale mandato si celano, infatti, problemi etici che chiamano in causa questioni etico-morali. A fronte di questa nuova situazione, si esige di uscire dalla visione esclusivamente tecnico-scientifica per abbracciare gli aspetti etico-morali dell'attività del medico. Se non viene mai messo in discussione che spetti al medico prospettare terapie e scenari di cura e che il medico sia vincolato, nel suo operare, al principio di beneficenza e non maleficenza, ciò che muta è il rapporto tra la prospettiva etica del medico e quella del paziente. Mentre nel modello paternalistico, l'etica medica è destinata a prevalere, sempre e comunque, nel modello che si afferma dagli anni Settanta in poi (modello liberale o modello informativo²³), è l'etica del paziente che prevale, attraverso il principio di autonomia. L'ultima parola sul se acconsentire o meno ai trattamenti/interventi proposti spetta al paziente, che manifesta la decisione con un'esplicita autorizzazione (consenso informato). Nell'attenersi al principio di beneficenza, il medico dovrà perseguire non un bene oggettivo ed astratto, ma un bene che si definisce sia tramite il rispetto dei criteri di appropriatezza terapeutica dei trattamenti offerti sia attraverso la valutazione della qualità della vita, che solo il paziente può contribuire a determinare.

Medico e paziente non sono più in una relazione asimmetrica dove il paziente è mero destinatario di decisioni, ma in una relazione, fondata sulla fiducia, dove il medico ha il ruolo centrale di fornire le informazioni necessarie al paziente per prendere una decisione consapevole e autonoma²⁴.

²² Si pensi agli straordinari risultati della scienza genetica, che con la terapia genica arriva ad incidere sugli elementi costitutivi dell'essere umano. Ma anche alle possibilità di controllo dei tempi e dei modi del morire con le tecniche di rianimazione e i trattamenti salvavita, oltre alla possibilità di intervenire sui processi procreativi con la procreazione assistita.

²³ Il Consiglio d'Europa descrive il modello informativo come segue «[t]his model vests the vast majority of decision-making power in the patient. The objective of clinical interactions "is for the doctor to provide the patient with all relevant information, for the patient to select the medical interventions he or she wants, and for the doctor to execute the selected interventions." Objective values are not assumed; rather, the patient's values and interests are taken as known or fixed to the patient but not to the doctor. The doctor's role is to provide facts to facilitate the patients making a decision that best matches their interests». Cfr. [https://www.coe.int/en/web/bioethics/theoretical-framework-of-the-doctor-patient-relationship#l%22123745941%22:\[1\]](https://www.coe.int/en/web/bioethics/theoretical-framework-of-the-doctor-patient-relationship#l%22123745941%22:[1]).

²⁴ Come è stato spiegato «respect for autonomy in health care relationships requires much more than avoiding deception and coercion. It requires an attempt to instill relevant understanding, to avoid forms of manipulation, and to respect persons' rights», T.L. BEAUCHAMP, J.F. CHILDRESS,

La letteratura bioetica sul tema ha sviluppato un'articolata riflessione, che, negli anni, ha portato a una chiara definizione del ruolo informativo del medico rispetto all'espressione del consenso informato da parte del paziente²⁵. In particolare, Beauchamp e Childress hanno, a suo tempo, distinto tre momenti del processo informativo e del conseguente momento autorizzativo²⁶, espresso dal consenso informato.

Il primo momento del processo prende il nome di *disclosure* (rivelare/comunicare l'informazione). La *disclosure* ha rappresentato, alle origini dell'elaborazione della dottrina del consenso informato, in particolare negli Stati Uniti, l'elemento su cui ci si è maggiormente concentrati. Per superare la storica reticenza del medico a informare il paziente, si considerava assolto il dovere di informazione da parte del medico con la rivelazione delle informazioni che riguardavano la situazione clinica del paziente.

Tuttavia, con il trascorrere del tempo, ci si è resi conto che, sebbene la rivelazione dell'informazione sia il requisito imprescindibile e propedeutico a qualsiasi decisione autonoma da parte del paziente, se tale rivelazione non è accompagnata dal controllo della comprensione dell'informazione da parte del destinatario, essa può risultare insufficiente ai fini dell'esercizio di una reale autonomia. Si è così introdotto il dovere per il medico di valutare il grado di comprensione da parte del paziente dell'informazione ricevuta (*understanding*). Una volta assicurata la comprensione, la fase autorizzativa, quella concretizzata dal consenso informato (*informed consent*), risulta di piena competenza del paziente, a cui dovrebbe essere lasciato un congruo lasso di tempo per ponderare la sua decisione, soprattutto laddove questa riguardi l'accettazione o il rifiuto di trattamenti e interventi altamente invasivi.

Le considerazioni che precedono, svolte sul piano del dibattito etico-morale, sono state opportunamente tradotte in regole giuridiche. Nel contesto europeo, il primo riferimento alla relazione di cura e agli strumenti per l'esercizio della libertà del paziente è rappresentato dalla già citata Convenzione di Oviedo, che dedica il Capitolo II al tema del consenso e delle disposizioni anticipate di trattamento. Segue la Carta dei diritti fondamentali dell'Unione europea che all'ar-

Principles of Biomedical Ethics, Oxford, Oxford University Press, 2013 p. 121. Il sopra richiamato rapporto dello Steering Committee for Human Rights in the field of Biomedicine and Health afferma che una caratteristica che definisce la medicina è la *healing relationship between clinicians and patients*, che coinvolge «both clinical interventions and information or services provided to patients for the sake of knowledge, empowerment or self-care», *ivi*, p. 38.

²⁵ In questa sede ci occupiamo solo del consenso informato e non delle disposizioni anticipate di trattamento, perché è con riferimento al primo che si pone la questione più pressante dell'uso dell'IA da parte del medico. Ciò non toglie che alcune riflessioni che verranno proposte potranno essere estese anche alle disposizioni anticipate di trattamento.

²⁶ T.L. BEAUCHAMP, J.F. CHILDRESS, *Principles of Biomedical Ethics*, p. 125.

articolo 3 prevede il consenso come fondamento per il rispetto dell'integrità della persona in ambito medico e biologico. In Italia si è arrivati nel 2017 a una esplicita disciplina giuridica del consenso informato e delle disposizioni anticipate di trattamento con la legge n. 219/2017.

Stando al condiviso orientamento interpretativo delle norme in materia di consenso da parte della giurisprudenza, della dottrina e della letteratura in materia di bioetica²⁷, il consenso informato è il criterio legittimante l'attività del medico. L'informazione, fornita dal medico al paziente, deve essere completa – non in assoluto, bensì rispetto al caso di specie²⁸ –, aggiornata, veritiera e comprensibile, ossia calibrata sul paziente che si ha di fronte e non su uno standard astratto. Il paziente può decidere di delegare l'informazione ad altri o di affidarsi completamente al medico, ma tale volontà deve essere adeguatamente documentata²⁹.

Come emerge dalle brevi riflessioni richiamate sul tema dell'informazione e del consenso informato, quali pilastri dell'attuale modello di relazione di cura, la dottrina del consenso informato, che, dalla metà del secolo scorso ad oggi, si è venuta definendo sempre più dettagliatamente, è uno degli aspetti portanti della relazione di cura ed è anche quello che rischia di essere travolto dall'introduzione dell'IA, come strumento di supporto delle decisioni mediche.

Nel paragrafo che segue daremo conto dei principali aspetti da tenere presenti quando la relazione di cura è mediata dall'IA.

3.2. La relazione di cura mediata dall'IA

Svariati sono i profili di criticità su cui ci si può soffermare nell'analizzare come potrebbe cambiare la relazione di cura mediata dall'IA. Per gli scopi di questo volume e per coerenza con l'analisi appena svolta sulla relazione di cura, l'attenzione sarà rivolta, per un verso, alle ricadute sulla professionalità, sulla com-

²⁷ Cfr. per tutti P. BORSELLINO, *Bioetica tra autonomia e diritto*, Zadig, Milano, 1999; P. BORSELLINO, *Consenso informato. Una riflessione filosofica-giuridica sul tema*, in *Salute e Società*, 2012, 3, pp. 17-41; P. BORSELLINO, *Bioetica tra 'moralì' e diritto*, Milano, Raffaello Cortina, 2018; P. BORSELLINO, *Norme in materia di consenso informato e di disposizioni anticipate di trattamento. Una conquista per i pazienti e per gli operatori sanitari*, in *Rivista Italiana di Cure Palliative*, 2018, XX, 1, pp. 1-10; P. BORSELLINO, *Consenso informato e autodeterminazione terapeutica nella legge sul testamento biologico*, in *Salute e Società*, 2018, 4, pp. 35-50.

Inoltre, T.L. BEAUCHAMP, J.F. CHILDRESS, *Principles of Biomedical Ethics*, cit.; C. FARALLI (a cura di), *Consenso informato in medicina. Aspetti etici e giuridici*, numero monografico di *Salute e Società*, 2012, 3, pp. 1-208; N.M. DOALLO, *El derecho al consentimiento informado del paciente. Una perspectiva iusfundamental*, Granada, Editorial Comares, 2021.

²⁸ Non devono essere trascurate informazioni, che nello specifico caso, se non conosciute inficiano l'autenticità del consenso, T.L. BEAUCHAMP, J.F. CHILDRESS, *Principles of Biomedical Ethics*, cit.

²⁹ La legge n. 219/2017 dichiara all'articolo 1, comma 3 che «il paziente può rifiutare in tutto o in parte di ricevere informazioni ovvero indicare i familiari o una persona di sua fiducia incaricati di riceverle».

petenza e sulla responsabilità del medico³⁰ e, per altro e connesso verso, alle ricadute sulla capacità decisionale del paziente. Si considererà lo scenario in cui l'IA è utilizzata a supporto e non in sostituzione della componente umana, essendo questa la situazione più prossima a noi nell'immediato futuro.

Partendo dal primo aspetto, valuteremo le criticità che l'IA solleva con riferimento al ruolo del medico, tenendo fermo lo schema comunicativo costituito dai momenti della *disclosure* e dell'*understanding*.

Molte problematiche che riguardano la figura del medico si possono, infatti, cogliere analizzando l'impatto dell'IA in queste due fasi della comunicazione dell'informazione al paziente e di spettanza del medico.

Come abbiamo sottolineato nel paragrafo che precede, il momento della rivelazione dell'informazione (*disclosure*) è propedeutico a qualsiasi scelta che il paziente potrà/dovrà operare. L'informazione, che il medico è tenuto a dare al paziente, deve contenere tutti quegli elementi, in assenza dei quali la decisione autonoma può considerarsi inficiata. Ad esempio, l'informazione che un determinato intervento ha una percentuale alta di rischio di determinare un'invalità permanente risulta essere cruciale per considerare il consenso valido, sia in termini morali che giuridici. Tuttavia, sebbene l'informazione debba essere completa nei sensi richiamati, ciò non significa che debba essere completa in assoluto. In altre parole, non si chiede al medico di riversare sul paziente tutta la conoscenza scientifica sul caso di specie, ma solo ciò che serve al paziente per una decisione consapevole, non manipolata, priva di coercizione di qualunque tipo. A fronte di questo stato dell'arte, come possiamo classificare l'informazione circa il supporto dato dall'IA al medico? È sempre un'informazione che il medico deve svelare? Vi sono circostanze in cui egli può non informare il paziente?

Per rispondere a queste domande occorre distinguere tra impiego a fini diagnostico-prognostici circa lo sviluppo di una patologia fisica o mentale e l'impiego a fini chirurgici.

Nel caso di utilizzo a supporto di decisioni diagnostiche, con fine implicitamente prescrittivo, se l'impiego dipende da un processo di innovazione della pratica clinica, ossia l'uso del sistema di IA non si è ancora dimostrato superiore agli standard adottati sino alla sua introduzione nella pratica clinica, andrà dimostrata la necessità di introdurre questo dispositivo innovativo, perché tale operazione si pone al confine tra pratica clinica e ricerca clinica. Bisognerà chiedersi se l'innovazione, che si vuole inserire, comporti rischi eticamente accettabili. A tal proposito, il medico deve poter valutare i criteri di validazione del sistema di IA che si vuole introdurre e comprenderne il potenziale virtuoso, ma anche i rischi. Nel valutare questi aspetti, il medico dovrebbe essere in grado di

³⁰ Cfr. G. PRAVETTONI, S. TRIBERTI, *Il Medico 4.0: come cambia la relazione di cura nell'era delle nuove tecnologie*, Milano, Edra, 2019.

comprendere la logica sottesa alla previsione che il sistema propone, se tale comprensione è rilevante per l'accuratezza della diagnosi che deve fare e se tale informazione rientra, quindi, tra quelle che vanno rivelate al paziente. Si pone, quindi, in relazione al momento della *disclosure* la pressante questione di avere a disposizione sistemi di IA addestrati con dati corretti e completi, in modo da evitare errori. Si pone, altresì, la necessità di mettere il medico in condizione di comprendere i sistemi di IA, venendo, ad esempio, coinvolto nel processo di progettazione e di comprensione dei *datasets* che ispirano la sua decisione finale³¹.

Questi passi sono necessari affinché il medico possa gestire al meglio sia la fase di *disclosure* dell'informazione al paziente, rimanendo fedele al dovere comunicativo e a quello di confidenzialità e riservatezza, sia la fase di comprensione dell'informazione, che costituisce il secondo momento del processo informativo, ossia l'*understanding*. Partendo dal presupposto che, quando l'informazione circa l'uso di sistemi di IA sia un presupposto ineludibile per una decisione autonoma da parte del paziente, è giocoforza dedurre che il controllo circa la comprensione sul ruolo giocato dall'IA nella decisione del medico sia un presupposto altrettanto inderogabile per una buona comunicazione, volta a garantire l'esercizio reale dell'autodeterminazione del paziente.

Molte delle riflessioni che precedono, sull'impiego dell'IA a fini diagnostico-terapeutico, sono adattabili anche al caso in cui l'IA sia utilizzata per il funzionamento di strumenti robotici a fini chirurgici³². Tuttavia, in questo ambito, ol-

³¹ Si muove in questa direzione la *Recommendation on the ethics of artificial intelligence* elaborata dalla *General Conference of the United Nations Educational, Scientific and Cultural Organization (UNESCO)*, nel 2021. Al punto 123 si legge:

123. Member States should pay particular attention in regulating prediction, detection and treatment solutions for health care in AI applications by:

- (a) ensuring oversight to minimize and mitigate bias;
- (b) ensuring that the professional, the patient, caregiver or service user is included as a “domain expert” in the team in all relevant steps when developing the algorithms;
- (c) paying due attention to privacy because of the potential need for being medically monitored and ensuring that all relevant national and international data protection requirements are met;
- (d) ensuring effective mechanisms so that those whose personal data is being analysed are aware of and provide informed consent for the use and analysis of their data, without preventing access to health care;
- (e) ensuring the human care and final decision of diagnosis and treatment are taken always by humans while acknowledging that AI systems can also assist in their work;
- (f) ensuring, where necessary, the review of AI systems by an ethical research committee prior to clinical use.

La Raccomandazione è accessibile a questo link <https://unesdoc.unesco.org/ark:/48223/pf0000380455>, ultimo accesso agosto 2022.

³² Un settore in grande sviluppo è quello della chirurgia spinale, dove la robotica supportata dall'IA promette risultati sorprendenti con meno fatica da parte dei chirurghi, cfr. J. RASOULI, J. SHAO, J. NEIFERT, M.P. STEINMETZ, E. BENZEL, T.E. MROZ, *Artificial Intelligence and Robotics in Spine Surgery*, in *Global Spine Journal*, 2021, 11, 4, pp. 556-564.

tre alla necessità di una specifica preparazione del medico sul funzionamento del sistema robotico ai fini della *disclosure*, è altrettanto importante che il medico sia in grado, nella fase di *understanding*, di valutare l'eventuale eccessiva fiducia o sfiducia riposta dal paziente nello strumento robotico e di aiutarlo a superare eventuali timori. L'ideologia antropomorfizzante e la sottesa idea che le macchine siano migliori degli umani, perché non influenzate dalla soggettività umana, ha radici profonde nella cultura contemporanea e il contesto medico non ne è esente. Questo convincimento può portare sia il medico sia il paziente a sopravvalutare le capacità del macchinico, sottostimando il rischio di errore. Dal punto di vista del medico, questo atteggiamento produce il c.d. fenomeno dell'*overreliance*, ossia dell'eccessiva fiducia nello strumento che lo assiste, che nel lungo termine può anche indurre un processo di *deskilling*, ossia di perdita delle proprie capacità mediche. Dal punto di vista del paziente, l'eccessiva fiducia riposta nell'IA può determinare un eccessivo affidamento alle sue potenzialità, ingenerando disillusione nel caso in cui qualcosa andasse storto.

D'altro canto, il paziente potrebbe avere una percezione falsata in senso negativo delle potenzialità e dei benefici dell'uso di un sistema robotico e insistere affinché l'intervento chirurgico venga effettuato in modo *tradizionale*. Anche in questo caso risulta fondamentale la preparazione del medico in merito alla prospettazione chiara dei vantaggi e dei limiti del sistema di IA proposto. Resta al momento non del tutto chiarito se il paziente possa pretendere che un intervento chirurgico non venga eseguito con l'assistente robotico, ma solo dalle mani del chirurgo. L'ipotesi è che, se l'assistente robotico rappresenta un'innovazione clinica non ancora validata nella prassi, il paziente non possa essere obbligato a sottoporsi a un intervento con caratteristiche sperimentali. Va, tuttavia, associata a questa riflessione, la previsione giuridicamente stabilita secondo la quale il medico non ha obblighi nei confronti di richieste del paziente che siano contrarie a norme di legge, alla deontologia professionale e alle buone pratiche clinico-assistenziali³³. Sul tema sarà certamente utile una presa di posizione normativa³⁴,

³³ Così recita l'articolo 1, comma 6 della legge n. 219/2017.

³⁴ Il Comité Consultatif National d'Éthique francese ha realizzato nel 2018 un rapporto, intitolato *Rapport de Synthèse. Bioéthique États Généraux*, in cui affronta diverse sfide scientifiche e tecnologiche tra cui quella dell'IA in medicina. Il rapporto sintetizza gli esiti della consultazione nazionale sui temi bioeticamente più scottanti. Con particolare riferimento al rapporto tra medico e paziente, il Comitato ritiene che il paziente debba essere messo a conoscenza dell'uso di sistemi di IA per poter esprimere un consenso libero e informato. Così si è espresso il Comitato «[i]l est par ailleurs demandé aux concepteurs de systèmes d'IA de respecter certains principes, pour que le «consentement éclairé» des personnes à utiliser des systèmes d'IA ne soit pas un leurre. Dans le cas d'un acte médical invasif ou pour un essai thérapeutique, le patient doit pouvoir être éclairé sur l'algorithme à l'origine de la proposition qui lui est faite par le médecin. Il lui faut au moins être assuré des informations ou raisonnements sur lesquels repose la réflexion médicale. Le concepteur doit veiller à ce que les résultats d'un système soient autant que possible interprétables et

che dia un quadro generale di principi a cui gli stati possano conformare le loro normative.

Nella fase di *disclosure* e in quella di *understanding* è opportuno che il medico sia in grado di chiarire, almeno a grandi linee, con il supporto di consulenti legali, chi è considerato responsabile in caso di errore durante l'intervento chirurgico. Questo aspetto è fondamentale per mantenere vivo il rapporto di fiducia con il paziente. In letteratura, infatti, viene utilizzata l'espressione "problem of many hands"³⁵ per indicare un'ulteriore problematicità derivante dalla difficoltà di attribuire responsabilità etica e giuridica quando a essere coinvolti sono diversi attori, i cui ruoli non sono chiaramente definiti. Sebbene questo sia un problema di politica pubblica e di diritto, appare evidente che il rapporto di fiducia tra medico e paziente esiga chiarezza anche con riferimento a questo complesso tema.

In aggiunta a queste riflessioni sulla relazione di cura mediata dall'IA, si vuole richiamare l'attenzione su un ulteriore fattore critico. Si tratta del rischio di esasperare, mediante il processo di *datificazione*³⁶, una visione astratta della malattia, che prescinde dalla persona in carne e ossa che ne soffre. Ciò si pone in palese contrasto con il processo, che la riflessione bioetica ha promosso negli ultimi settant'anni, volto a mettere il paziente, inteso come persona con i suoi bisogni e le sue credenze, al centro della relazione di cura. Questo rischio è accentuato dalla trasformazione del ruolo del medico, che nel nuovo scenario mediato dall'IA, si configura sempre più come un medico non tanto specialista quanto come tecnico specializzato, che si occupa del corpo macchina, senza occuparsi della sofferenza e dell'assistenza al paziente.

4. Neurotecnologie e potenziamento umano: le nuove frontiere della medicina

Le applicazioni di IA risultano molto promettenti in campo medico quando associate alle neurotecnologie. Queste ultime sono tecnologie che consentono di stabilire una connessione con il cervello attraverso la quale si rende possibile mo-

explicables pour les utilisateurs», ivi, p. 106. Ha fatto seguito al rapporto una modifica delle Legge sulla Bioetica, Loi n. 2021-1017 relative à la bioéthique. La legge ha introdotto all'articolo 17 un nuovo articolo per il codice di sanità pubblica che così recita «Art. L. 4001-3.-I.-Le professionnel de santé qui décide d'utiliser, pour un acte de prévention, de diagnostic ou de soin, un dispositif médical comportant un traitement de données algorithmique dont l'apprentissage a été réalisé à partir de données massives s'assure que la personne concernée en a été informée et qu'elle est, le cas échéant, avertie de l'interprétation qui en résulte».

³⁵D. SCHIFF, J. BORENSTEIN, *How Should Clinicians Communicate With Patients About The Roles Of Artificially Intelligent Team Members?*, in *AMA Journal of Ethics*, 2019, 21, pp. E 138-145.

³⁶F. CABITZA, R. RASOINI, G.F. GENSINI, *Unintended consequences of machine learning in medicine*, *Jama*, 2017, 318, 6, pp. 517-518.

nitorare e registrare³⁷ e/o alterare l'attività neuronale³⁸. Vi sono tre aree in cui si è soliti suddividere le neurotecnologie, che consentono di intervenire sull'attività neuronale (*open-loop or closed-loop neurotechnologies*³⁹), e sono le seguenti: 1) tecnologie di neurostimolazione; 2) tecnologie di neuroprotesi; 3) interfacce uomo-macchina (*brain-machine/computer interfaces*, *BMIs* o *BCIs*). La prima categoria include ogni dispositivo che si avvale di un'interfaccia neuronale per stimolare parti del sistema nervoso centrale, periferico e autonomo. Esse possono essere tecnologie *open-loop* o *closed-loop*. Un esempio di questa seconda categoria è la terapia di neuromodulazione per i pazienti affetti da Parkinson. In tal caso, i sensori applicati al paziente adattano la stimolazione neuronale per limitare il tremore associato alla patologia.

Le tecnologie di neuroprotesi sono dispositivi utili a sostituire funzioni cognitive, motorie o sensoriali. Un esempio di questa tipologia sono gli impianti cocleari, che consentono alle persone affette da perdita dell'udito di interpretare i segnali attraverso la stimolazione dei neuroni uditivi nel tronco encefalico.

Infine, le interfacce uomo-macchina (*brain-machine interfaces*, *BMIs*) sono dispositivi che collegano direttamente il cervello a un dispositivo esterno. Possono avere funzione unidirezionale, consentendo all'utilizzatore di controllare un dispositivo esterno o di ricevere degli stimoli sensori da un dispositivo esterno, ma possono anche essere bidirezionali e consentire sia l'attività di codificazione che di decodifica. Sono esempi di queste interfacce uomo-macchina quelle applicate ai pazienti affetti da paresi. Tali interfacce consentono di leggere e decodificare i segnali cerebrali ai fini di controllare arti robotici e di ritrascrivere nel cervello, attraverso la stimolazione, la sensazione sensoriale naturale. Le interfacce uomo-macchina sono utilizzate in campo terapeutico per trattare, attraverso la stimolazione cerebrale profonda (DBS), malattie come la depressione. Sono, quindi, in grado di agire direttamente sulle condizioni delle nostre emozioni. Queste neurotecnologie sono altresì utilizzate per trattare deficit cognitivi, causati da malattie come l'Alzheimer, migliorando le prestazioni mnemoniche dei pazienti, anche attraverso il recupero di c.d. *memories flashback*.

In sintesi, in ambito medico molti sono gli straordinari risultati già ottenuti o promessi dall'uso di queste tecnologie per curare patologie fisiche e mentali, migliorando la qualità di vita dei pazienti⁴⁰.

³⁷ Il più utilizzato è l'elettroencefalogramma (EEG) e la stimolazione magnetica transcranica (TMS).

³⁸ Cfr. M. IENCA, *Common Human Rights Challenges Raised by Different Applications of Neurotechnologies in the Biomedical Fields*, Report Commissioned by the Committee on Bioethics (DH-BIO) of the Council of Europe, 2021, accessibile a questo link <https://www.coe.int/en/web/bioethics/round-table-on-the-human-rights-issues-raised-by-the-applications-of-neurotechnologies>.

³⁹ Le tecnologie *open-loop* offrono una stimolazione senza feedback, mentre le *closed-loop* consentono di adattare i parametri di stimolazione sulla base del feedback che ricevono.

⁴⁰ Un potenziale uso di queste tecnologie può essere ipotizzato nell'ambito delle decisioni di

Queste stesse tecnologie hanno però un potenziale utilizzo che esula dalla stretta terapia. Nel 2009 il *Panel for the Future of Science and Technology* (STOA) pubblica il primo ampio studio sul tema del potenziamento umano⁴¹, in cui un intero paragrafo è dedicato alle tecnologie di stimolazione cerebrale profonda e alle loro prospettive in campo terapeutico e non terapeutico e alle correlate criticità etiche. Ciò dimostra come il tema sia ormai oggetto di riflessione, da più di un decennio, nel contesto istituzionale europeo. Per comprendere i risvolti degli sviluppi neurotecnologici correlati all'IA a fini non terapeutici, appare opportuno dare, preliminarmente, qualche coordinata per orientarsi nel complesso dibattito sul potenziamento umano.

4.1. *Il biopotenziamento in breve*

Con il termine potenziamento umano (*human enhancement*) si indicano gli interventi sul corpo o sul cervello umano resi possibili dagli sviluppi in ambito scientifico e tecnologico, volti a migliorare e/o aumentare capacità e abilità umane preesistenti o a crearne di nuove⁴². Gli interventi potenzianti possono essere di tipo terapeutico, quando il trattamento della patologia di un paziente gli consente di ottenere risultati di *performance* che superano gli standard tipici di salute, ad esempio quando, grazie a una protesi, un soggetto può avere prestazioni fisiche migliori di un normodotato (caso Pistorius). Ma in genere, quando si parla di potenziamento umano si fa riferimento agli interventi su soggetti sani senza scopo terapeutico.

A livello internazionale, il dibattito sul potenziamento umano così inteso è iniziato negli anni Sessanta del secolo scorso negli Stati Uniti, con l'interessamento

fine vita. A mezzo di queste interfacce potrebbe essere possibile per pazienti in grado di decidere autonomamente, ma non in grado di attivare fisicamente alcun dispositivo, di essere aiutati nell'esecuzione dell'atto finale nel caso di suicidio medicalmente assistito.

⁴¹ PANEL FOR THE FUTURE OF SCIENCE AND TECHNOLOGY (STOA), *Human Enhancement Study*, 2009, accessibile a questo link [https://www.europarl.europa.eu/stoa/en/document/IPOL-JOIN_ET\(2009\)417483](https://www.europarl.europa.eu/stoa/en/document/IPOL-JOIN_ET(2009)417483).

⁴² Allen Buchanan definisce il potenziamento umano come «a deliberate intervention, applying biomedical science, which aims to improve an existing capacity that most or all normal human beings typically have, or to create a new capacity, by acting directly on the body or brain», cfr. BUCHANAN, *Beyond humanity*, Oxford, Oxford University Press, p. 11. Il Panel for the Future of Science and Technology definisce il potenziamento umano nel sopra citato studio come «any modification aimed at improving individual human performance and brought about by science-based or technology-based interventions in the human body», p. 6. Peter Conrad ha individuato anche un uso delle tecniche che vanno sotto l'etichetta di biopotenziamento per ottenere una "normalizzazione" o "standardizzazione" di una data condizione, ad esempio, l'uso dell'ormone della crescita, cfr. P. CONRAD, *The Medicalization of Society. On Transformation of Human Conditions into Treatable Disorders*, Baltimore, John Hopkins University Press, 2007, p. 87. Questa accezione di biopotenziamento non è più contemplata nella letteratura recente sul tema.

mento di alcuni sociologi⁴³, che lo hanno inserito nella più ampia riflessione attorno alla medicalizzazione della società⁴⁴. Nella prima fase del dibattito, la riflessione sul potenziamento ha riguardato per lo più l'utilizzo di farmaci – prodotti per uso terapeutico – a fini non terapeutici, bensì potenziativi. Si tratta di farmaci prodotti per trattare stati patologici, impiegati invece per aumentare le capacità cognitive – mnemoniche e di concentrazione – di soggetti sani o per migliorare le loro prestazioni fisiche (doping nello sport)⁴⁵. Tuttavia, prima sul piano accademico e poi anche a livello istituzionale, il dibattito sul potenziamento umano si è esteso e oggi ricomprende considerazioni sulle criticità dell'impiego di dispositivi e farmaci per trattare non solo aspetti correlati all'attività fisica o alle prestazioni cognitive, ma si spinge fino a proporre di intervenire sugli aspetti morali degli individui. Si tratta del c.d. controverso progetto di biopotenziamento morale⁴⁶, volto a incidere sulle qualità morali degli individui con il fine di creare una società di “buoni”.

Il potenziamento umano, in tutte le sue sfaccettature, solleva diversi quesiti: filosofici, su come vada intesa la natura umana, malleabile e modificabile a piacere o limitata e statica; etici, sull'accettabilità e giustificabilità del miglioramento e potenziamento artificiale dell'uomo, sui rischi accettabili, nonché sulla valutazione dell'assunto culturale se ciò che è meglio è anche necessariamente buono e, quindi, da perseguire⁴⁷; giuridici, in merito alle risposte che il diritto

⁴³ Cfr. J. PITTS, *Social Control: The Concept*, in D. SILLS (eds.), *International Encyclopaedia of Social Sciences*, New York, Macmillan, 1968, vol. 14; I.K. ZOLA, *Medicine as an Institution of Social Control*, in P. CONRAD (eds.), *The Sociology of Health and Illness: Critical Perspectives*, New York, Worth, 1972, pp. 497-506.

⁴⁴ Per tutti cfr. P. CONRAD, *The Medicalization of Society*, cit. L'autore ha richiamato la questione definitoria circa la medicalizzazione della società, sottolineando che dire di un evento o processo naturale (sessualità, nascita, morte ecc.) che è medicalizzato, significa descriverlo e definirlo in termini medici e, quindi, trattarlo con interventi di tipo medico.

⁴⁵ Cfr. sul tema del potenziamento umano nell'ambito sportivo S. SALARDI, *Lo sport come diritto umano nell'era del post-umano*, Torino, Giappichelli, 2019. Per uno sguardo generale, L. PALAZZANI, *Potenziamento umano: tecnoscienza, etica e diritto*, Torino, Giappichelli, 2015.

⁴⁶ Il tema è stato inizialmente introdotto da I. PERSSON, J. SAVULESCU, *The Perils of Cognitive Enhancement and the Urgent Imperative to Enhance the Moral Character of Humanity*, in *Journal of Applied Philosophy*, 2008, 25, 3, pp. 162-177. Per un approfondimento sul tema a S. SALARDI, *When the 'age of rights' meets the 'age of technology'*, in A. D'ALOIA, M. ERRIGO (a cura di), *Neuroscience and Law. Complicated Crossings and New Perspectives*, Cham, Springer, 2020, pp. 239-255; S. SALARDI, *The 'Project of Moral Bioenhancement' in the European legal system. Ethically controversial and legally highly questionable*, in *Rivista di Filosofia del Diritto – Journal of Philosophy of Law*, 2018 VII, 2, pp. 241-260; S. SALARDI, *Destined to be superhuman. Moral bioenhancement and its legal viability*, in *Biolaw Journal*, 2017, 3, pp. 87-101; M. GALLETTI, *Il biopotenziamento morale e il bilanciamento tra libertà e bene*, in S. SALARDI, M. SAPORITI, M. VETIS ZAGANELLI (a cura di), *Diritti umani e tecnologie morali. Una prospettiva comparata tra Italia e Brasile*, pp. 87-100, cit.

⁴⁷ Cfr. N. LEVY, *Neuroethics. Ethics and the Science of the Mind*, in U. SCHÜKLENK, P. SINGER (eds.), *Bioethics. An Anthology*, New Jersey, Wiley Blackwell, 2022, pp. 861-865.

deve apprestare, individuando un bilanciamento di interessi il più possibile equo⁴⁸ e se per fare ciò sia meglio utilizzare principi generali già esistenti o elaborarne di nuovi (ad esempio, il dibattito sui neurodiritti).

Dal punto di vista filosofico il potenziamento umano è guardato con sfavore dai bioconservatori e promosso dagli esponenti del transumanesimo⁴⁹, «tenacemente impegnati, i primi, a restaurare i diritti della natura. Custodi, gli altri, di una nuova libertà, quella appunto di usare senza limiti l'inedito potere di cui siamo investiti»⁵⁰. Il potenziamento umano è una tematica che si pone al crocevia tra diverse questioni, al cui centro vi è la tenuta della "sfera intima" del soggetto e dei principi che la proteggono a livello normativo. Il potenziamento umano mette, infatti, in gioco i diritti della persona in diversi ambiti in cui questa si trova a operare: 1) diritti dei consumatori a fronte del crescente commercio di dispositivi direttamente venduti al consumatore (*direct-to-consumer devices*) per l'autopotenziamento ricreativo, ma non solo; 2) diritti della persona nel *military enhancement*⁵¹, che è una pratica spesso imposta e non scelta; 3) diritti della persona e volto della medicina.

Al di là degli eccessivi entusiasmi transumanisti e del catastrofismo di alcune posizioni bioconservatrici⁵², è innegabile che il potenziamento umano è ormai parte della medicina. A tal proposito basta richiamare l'articolo 76 del Codice di deontologia medica italiano del 2014, esplicitamente dedicato alla medicina potenziativa⁵³. Nonostante il potenziamento umano sia già parte della deontologia

⁴⁸ In generale in materia di sviluppi neuroscientifici, sul ruolo del diritto cfr. A. FARANO, *Neuroscienze e diritto: un primo bilancio*, in S. SALARDI, M. SAPORITI (a cura di), *Le tecnologie 'moralì' emergenti e le sfide etico-giuridiche delle nuove soggettività*, Torino, Giappichelli, 2020, pp. 42-51.

⁴⁹ Questa corrente di pensiero sostiene la desiderabilità di un passaggio dalla condizione umana a quella post-umana ai fini, per un verso, di un miglioramento delle capacità fisiche e intellettuali dell'uomo e, per altro verso, dell'eliminazione di alcune limitazioni come l'invecchiamento. Nella visione transumanista la garanzia di accesso alla tecnologia e la sua disponibilità a tutti consentirà il passaggio dall'uomo al *cyborg*, agevolando la transizione dalla condizione umana a quella post-umana, che è, in questa prospettiva, l'inevitabile approdo dell'evoluzione umana.

⁵⁰ S. RODOTÀ, *Il diritto di avere diritti*, Roma-Bari, Laterza, 2012, p. 355.

⁵¹ U. RUFFOLO, A. AMIDEI, *Intelligenza Artificiale, human enhancement e diritti della persona*, in U. RUFFOLO (a cura di), *Intelligenza Artificiale-Il diritto, i diritti, l'etica*, Milano, Giuffrè, 2020, pp 179-211.

⁵² Secondo, ad esempio, Paolo Benanti, il potenziamento umano non terapeutico rappresenta un pericolo per la medicina, in quanto incide sull'identità e sul ruolo del medico, sui valori dell'etica medica e trasforma la medicina da contesto di cura a quello del miglioramento con il correlato passaggio da paziente a cliente, privando la categoria della cura del suo presunto significato intrinsecamente essenziale. Cfr. P. BENANTI, *The Cyborg: Corpo e corporeità nell'epoca del post-umano*, Assisi, Cittadella, 2012.

⁵³ L'articolo così recita al comma 1 «Il medico, quando gli siano richiesti interventi medici finalizzati al potenziamento delle fisiologiche capacità psico-fisiche dell'individuo, opera, sia

medica, non vi è una regolamentazione giuridica espressamente dedicata al tema. Tuttavia, la crescente disponibilità di mezzi per il potenziamento di soggetti sani esige una risposta giuridica, soprattutto se si pensa alle possibilità offerte dalle neurotecnologie.

Molte prospettive aperte dalle terapie a mezzo di neurotecnologie in ambito medico potranno, nel tempo, venire estese al di fuori di questo contesto. In alcuni casi, è necessario l'intervento medico – ad esempio per impiantare chip nel cervello –, ma in altri casi, tale intervento non è necessario, trattandosi di dispositivi indossabili (*wearables*), alcuni dei quali già in commercio. Rispetto ai due indicati scenari si proporranno di seguito alcune brevi considerazioni etico-giuridiche.

4.2. *Profili etico-giuridici del potenziamento neurotecnologico*

Come valutare le prospettive aperte dalle neurotecnologie per un potenziamento non terapeutico⁵⁴? Le valutazioni sono diverse a seconda che le neurotecnologie richiedano interventi medici o al contrario siano direttamente vendute al consumatore. La valutazione deve altresì tenere conto del grado di invasività della tecnologia in questione, che si muove entro i confini di due situazioni estreme: da una parte, interventi chirurgici reversibili (es. impianto di chip nel cervello) o irreversibili (terapia genica); dall'altra parte, dispositivi neurotecnologici indossabili, usati direttamente dai consumatori.

Consideriamo dapprima le forme di potenziamento neurotecnologico invasivo, che richiedono, quindi, l'intervento di un medico.

Rispetto a queste neurotecnologie si pongono, in primo luogo, questioni relative all'accettabilità di questo tipo di potenziamento all'interno della medicina. Come ricordato nel paragrafo precedente, le posizioni sul tema sono molto articolate, ma non si può negare che la deontologia abbia già preso in considerazione il fenomeno del potenziamento umano, fornendo indicazioni al medico per garantire un alto livello di protezione del paziente.

In secondo luogo, va ricordato che laddove vi sia l'intervento del medico per forme di potenziamento invasive, questa attività, rientrando nei confini della relazione di cura, è assoggettata a tutte le norme che garantiscono diritti e prevedono doveri in quello specifico ambito. Viene garantito così il fondamentale di-

nella fase di ricerca che nella pratica professionale, secondo principi di precauzione, proporzionalità e rispetto dell'autodeterminazione della persona, acquisendo il consenso informato in forma scritta».

⁵⁴Per uno sguardo alle implicazioni etiche delle neurotecnologie cfr. INTERNATIONAL BIOETHICS COMMITTEE (IBC), *Report of the International Bioethics Committee of UNESCO (IBC) on the Ethical Issues of Neurotechnology*, 2021, accessibile a questo link <https://unesdoc.unesco.org/ark:/48223/pf0000378724>.

ritto all'informazione corretta e al controllo di tale informazione per un esercizio di autonomia reale da parte del soggetto che richiede l'intervento. Dentro la cornice normativa della relazione di cura, la garanzia di questo imprescindibile diritto trova la sua massima espressione. È pertanto auspicabile che la medicina si faccia carico esplicito di queste forme di potenziamento, che, nella relazione di cura, basata sulla fiducia e sulle autonomie di medico e paziente, possono trovare un baluardo contro abusi e manipolazioni.

Sul piano del diritto, forme di potenziamento neurotecnologico, che mettono in connessione il cervello umano con dispositivi esterni, sono state valutate alla luce della cornice dei diritti fondamentali. Posto che non esiste, nel contesto europeo, un diritto esplicito al potenziamento umano, esiste però un diritto al libero sviluppo della propria personalità, sulla base del quale si potrebbe giustificare il potenziamento anche nelle forme invasive. Tale possibilità di estensione è tuttavia ancora da valutare e non trova un consenso nel dibattito etico europeo. Ma prescindendo ora da questa specifica problematica, ciò che è oggi in discussione è la protezione della nostra personalità/identità/integrità mentale rispetto alla possibilità, che attraverso le neurotecnologie (interfacce uomo-macchina), altri possano accedere ai nostri pensieri, ricordi, credenze ecc. La riflessione giuridica in questo campo si muove in due principali direzioni. Vi sono autori che ritengono che i diritti fondamentali enucleati nelle costituzioni siano sufficienti a proteggere da queste inedite forme di invasione della sfera privatissima dell'individuo, potendo contare sull'interpretazione estensiva da parte delle Corti. Altri autori sostengono l'improrogabilità della previsione di nuovi diritti, c.d. *neurorights* o neurodiritti⁵⁵. La battaglia per i neurodiritti è approdata, in alcuni ordinamenti fino a toccare la riflessione costituzionale. È il caso del Cile, dove sono state proposte modifiche alla Costituzione per introdurre la protezione dell'integrità mentale all'articolo 19, comma 1.

Se dalle considerazioni sulle neurotecnologie, che richiedono un intervento medico, ci spostiamo a considerare quelle direttamente offerte al consumatore, ossia dispositivi direttamente utilizzabili senza mediazione esperta, ci avvediamo che in questo specifico caso, mancando la possibilità di un rapporto di fiducia con un esperto, molte delle questioni etiche e giuridiche, che inserite nella relazione di cura trovano una definizione regolativa, in questo specifico campo re-

⁵⁵Per una ricognizione sulla letteratura in materia di neurodiritti: M. IENCA, R. ANDORNO, *Towards new human rights in the age of neuroscience and neurotechnology*, in *Life Sciences, Society and Policy*, 2017, 13, 1, pp. 1-28; P. SOMMAGGIO, *Neurodiritti: tra neuroscienze e neurotecnologie*, in S. SALARDI, M. SAPORITI, M. VETIS ZAGANELLI (a cura di), *Diritti umani e tecnologie morali. Una prospettiva comparata tra Italia e Brasile*, cit. pp. 159-174; F. STOCCHI, *Presupposti etico-filosofici, casi studio e bilanciamento di interessi nella regolamentazione neurotecnologica*, in *Federalismi.it*, 2021, 29, pp. 89-120. T. ISTACE, *Neurorights: the debate about new legal safeguards to protect the mind*, in *Issues in law and medicine*, 2022, 37, 1, pp. 95-114; M. IENCA, *On Neurorights*, in *Front Hum Neurosci*, 2021, 15, pp. 1-11.

stano invece inevase. In particolare, restano ancora da definire le seguenti criticità: 1) come rendere disponibile ai consumatori di neurotecnologie – a scopo di intrattenimento o rilassamento – adeguate informazioni sui potenziali rischi che il loro uso non supervisionato può indurre?; 2) come garantire che all’informazione segua un controllo della sua comprensione ai fini di una scelta realmente autonoma?; 3) stante l’enorme potenziale del mercato dei c.d. *brain data*⁵⁶, come proteggere i consumatori da logiche economiche che influenzano i mercati e a volte gli stessi decisori istituzionali in merito alla messa in commercio di certi dispositivi?

Le possibili risposte ai citati quesiti possono prendere spunto dalle raccomandazioni che proponiamo di seguito, ma necessiteranno di un’approfondita analisi sul piano giuridico per la predisposizione di strumenti di tutela adeguati da parte del decisore politico.

Le raccomandazioni generali proposte, senza presunzione di esaustività, per impostare il discorso giuridico possono essere così riassunte:

- prevedere processi di comunicazione di una corretta informazione sui rischi e benefici delle neurotecnologie usate a fini potenzianti e migliorativi attraverso canali istituzionali, dove il pubblico può interagire direttamente con gli esperti, pubblicizzando adeguatamente tali canali;
- garantire la possibilità di contattare un medico competente in caso di possibili effetti collaterali derivanti dall’uso delle neurotecnologie indossabili;
- garantire un’informazione veritiera e trasparente circa gli interessi commerciali che si celano dietro la spinta per una diffusione su larga scala di queste tecnologie;
- progettare e costruire in modo etico le neurotecnologie con cui è possibile accedere ai dati e ai pensieri immagazzinati nel nostro cervello, esponendo la persona a rischi di manipolazione della personalità.

4.3. *Neurotecnologie e ideologia antropomorfizzante: alcune riflessioni conclusive*

Le sopra richiamate raccomandazioni sono utili ai fini di superare la dialettica degli opposti e l’ideologia antropomorfizzante, che trovano nel connubio tra IA e neurotecnologie un importante alleato.

Possiamo per certi versi considerare le attuali interfacce uomo-macchina come gli eredi della concezione del *Cyborg* maturata nel secolo scorso, proprio nel contesto medico.

⁵⁶ Cfr. A.P. KARANASIOU, *On Being Trans-Human: Commercial BCIS and the Quest for Autonomy*, in W. BARFIELD (eds.) *The Cambridge Handbook of the Law of Algorithms*, Cambridge, Cambridge University Press, 2021, pp. 757-774.

Il termine nasce, infatti, nell'ambito biomedico, in particolare nel contesto delle ricerche spaziali della NASA tra la fine degli anni Cinquanta e l'inizio degli anni Sessanta del secolo scorso, coniato da due medici, Clynes e Kline⁵⁷. I due medici sperimentano un congegno meccanico da inserire nel corpo degli astronauti per iniettare sostanze biochimicamente attive, utili a superare i limiti fisici e psichici derivanti dalla permanenza nello spazio in assenza di gravità.

Il *cyborg* è oggi considerato dai transumanisti la fase di passaggio dalla condizione umana, limitata e imperfetta, alla condizione post-umana piena, in cui la dimensione corporea non sarà più necessaria. Il fenomeno del *Metaverso*⁵⁸ è esemplificativo dell'evoluzione rapida di questa concezione e dell'enorme spinta commerciale che si cela dietro la filosofia transumanista. Si ritiene che un giorno sarà possibile fare l'*upload* del nostro cervello dentro questa realtà virtuale che ci garantirà vita eterna. Alla base di questa visione dell'umano vi è l'assunto, per certi versi paradossale, che l'uomo, rielaborato dal macchinico da lui creato, sarà una versione migliorata di sé stesso. Partendo da questo assunto, segue che l'essere umano deve puntare a utilizzare la tecnologia a disposizione per raggiungere questo miglioramento, sino alle estreme conseguenze, per un verso, di rinunciare alla sua corporeità e, per altro e connesso verso, di rendere la sua mente aperta e disponibile a tutti.

Il processo di trasformazione dell'uomo in macchina richiederà per un certo lasso di tempo la collaborazione della medicina, ecco perché i transumanisti spingono senza tregua affinché quest'ultimo contesto incorpori, senza troppi problemi, la dimensione del potenziamento.

La proposta transumanista ha un carattere profondamente prescrittivo perché pretende di stabilire, artatamente, ciò che è buono in assoluto e pertanto da perseguire. Tale proposta individua e prescrive al contempo le proprietà da abbandonare e quelle da sviluppare, dando per scontato che tutti condividano la selezione di tali proprietà.

In particolare, facendo leva sulle primordiali paure dell'umanità – la morte e l'invecchiamento – il transumanesimo si pone come alternativa salvifica attraverso la messa a disposizione della tecnologia. Senza che ce ne rendiamo conto viene implicitamente prescritta la vita perfetta che pochi hanno deciso per noi, senza mai consultarci. Come è stato osservato «con il transumanesimo, emerge un nuovo paradigma religioso: quello dell'Uomo-Dio. Non è più la rinuncia dell'ateo, che si considera solo nell'universo, bensì l'orgogliosa af-

⁵⁷ Cfr. M.E. CLYNES, N.S. KLINE, *Cyborgs in Space*, in G. HABLES GRAY, H.J. FIGUEROA-SARRIERA, S. MENTOR (eds.), *The Cyborg Handbook*, New York, Routledge, 1995, pp. 30-31.

⁵⁸ È un fenomeno complesso che si riferisce a un universo digitale immersivo, a cui si accede con diversi strumenti tecnologici già in uso per la realtà virtuale e la realtà aumentata.

fermazione di tutto ciò che l'uomo può fare, incluso creare vita e ricreare sé stesso»⁵⁹.

Questa posizione prescrittiva mal si adatta a essere sussunta sotto il quadro valoriale sotteso alla cornice giuridica dei diritti fondamentali, concepita come limite ai poteri forti e, al contempo, come garanzia dei diritti dei più deboli. Questa prospettiva mal si concilia con la filosofia transumanista che promuove un'innovazione senza regole, o meglio con le sole regole del libero mercato e che non si interessa di costruire l'autonomia e l'autodeterminazione individuale attraverso seri percorsi istituzionali, culturali e anche giuridici.

Tuttavia, se vogliamo essere protagonisti del processo innovativo in atto e soprattutto se vogliamo essere attori che controllano l'inedito potere a disposizione, dobbiamo essere disposti a riconoscere che l'etica e il diritto sono i più importanti strumenti a disposizione per contrastare il crescente potere di pochi su molti. Sono gli strumenti irrinunciabili per riaffermare cioè che i diritti fondamentali costituiscono la *sfera dell'indecidibile*⁶⁰ e si sottraggono alle logiche di mercato, per riportare al centro la dignità della persona costituzionalizzata.

⁵⁹Cfr. L. ALEXANDRE, *La guerra delle intelligenze. Intelligenza artificiale contro Intelligenza umana*, Torino, EDT, 2018, p. 247.

⁶⁰L. FERRAJOLI, *Principia Juris. Teoria del diritto e della democrazia*, Roma-Napoli, Laterza, 2007.

BIBLIOGRAFIA

- ABBAGNANO N., *Dizionario di Filosofia*, Torino, Utet, 2013.
- ALEXANDRE L., *La guerra delle intelligenze. Intelligenza artificiale contro intelligenza umana*, Torino, EDT, 2018.
- ALFANO M., CHRISTEN M., *A framework for understanding and evaluating moral technologies*, in S. SALARDI, M. SAPORITI (a cura di), *Emerging 'moral' technologies and the ethical-legal challenges of new subjectivities*, Torino, Giappichelli, 2020, pp. 1-21.
- AMATO S., *Biodiritto 4.0. Intelligenza artificiale e nuove tecnologie*, Torino, Giappichelli, 2020.
- AMATO MANGIAMELI A.C., *Algoritmi e big data*, in *Rivista di Filosofia del Diritto*, 2019, VII, 1, pp. 107-124.
- ANGELETTI L.R., GAZZANIGA V., *Storia, filosofia ed etica generale della medicina*, Milano, Elsevier, 2012.
- ANSUÁTEGUI ROIG F.J., *Norme, Giudici, Stato Costituzionale. Frammenti di un Modello Giuridico*, Torino, Giappichelli, 2020.
- AUSTIN J.L., *A Plea of Excuses*, in *Proceedings of the Aristotelian Society*, 1956-1957, 57, pp. 1-30.
- BALAZKA D., RODIGHIERO D., *Big Data and the Little Big Bang: An Epistemological (R)evolution*, in *Frontiers in Big Data*, 2020, 3, 31, pp. 1-10.
- BALISTRERI M., CAPRANICO G., GALLETTI M., ZULLO S., *Biotecnologie e modificazioni genetiche*, Bologna, Il Mulino, 2020.
- BARASSI V., *I figli dell'algoritmo. Sorvegliati, tracciati, profilati dalla nascita*, Roma, LUISS University Press, 2021.
- BARBALET J., *The experience of Trust: Its Content and Basis*, in M. SAZAKI (eds.), *Trust in Contemporary Society*, USA, BRILL, 2019, pp. 11-30.
- BARBERIS M., *Filosofia del diritto*, Torino, Giappichelli, 2008.
- BARFIELD W., BARFIELD J., *An Introduction to Law and Algorithms*, in W. BARFIELD (eds.) *The Cambridge Handbook of The Law of Algorithms*, Cambridge, Cambridge University Press, 2021, pp. 3-15.
- BARLUCCHI C., *Causalità come congiunzione probabile*, in *Quaderni di Sociologia*, 2000, 22, pp. 81-104.
- BEAUCHAMP T.L., CHILDRESS J.F., *Principles of Biomedical Ethics*, Oxford, Oxford University Press, 2013.
- BENANTI P., *The cyborg: corpo e corporeità nell'epoca del post-umano*, Assisi, Cittadella editrice, 2016.

- BERLIN I., *Due concetti di libertà*, in H. HARDY (a cura di), *Libertà*, Milano, Feltrinelli, 2005.
- BERNARD C., *Introduzione allo studio della medicina sperimentale*, Padova, Piccin, 1994.
- BERTAGNA G., CERETTI A., MAZZUCCATO C. (a cura di), *Il libro dell'incontro*, Milano, Il Saggiatore, 2015.
- BOBBIO N., *La natura delle cose*, in ID., *Giusnaturalismo e Positivismo giuridico*, Edizioni Comunità, Milano, 1972, pp. 197-212.
- BOBBIO N., *Norma*, in ID., *Contributi ad un Dizionario Giuridico*, Torino, Giappichelli, 1994, pp. 177-214, riedizione della voce *Norma*, in *Noviss. Dig. it.*, 1964.
- BOBBIO N., *L'età dei diritti*, Torino, Einaudi, 1997.
- BOBBIO N., *Giusnaturalismo e positivismo giuridico*, Roma-Bari, Laterza, 2011.
- BORSELLINO P., *Bioetica tra autonomia e diritto*, Milano, Zadig, 1999.
- BORSELLINO P., *Morale, diritto e positivismo giuridico nel pensiero di Uberto Scarpelli*, in *Bioetica tra autonomia e diritto*, Milano, Zadig, 1999, pp. 213-276.
- BORSELLINO P., *Consenso informato. Una riflessione filosofico-giuridica sul tema*, in *Salute e Società*, 2012, 3, pp. 17-41.
- BORSELLINO P., *Bioetica tra "moralì" e diritto*, Milano, Raffaello Cortina, 2018.
- BORSELLINO P., *Consenso informato e autodeterminazione terapeutica nella legge sul testamento biologico*, in *Salute e Società*, 2018, 4, pp. 35-50.
- BORSELLINO P., *Norme in materia di consenso informato e di disposizioni anticipate di trattamento. Una conquista per i pazienti e per gli operatori sanitari*, in *Rivista Italiana di Cure Palliative*, 2018, XX, 1, pp. 1-10.
- BORSELLINO P., *Storicità del diritto e filosofia di orientamento analitico-linguistico. Quale rapporto?* in A. BALLARINI (a cura di), *La storicità del diritto. Esistenza materiale, filosofia, ermeneutica*, Torino, Giappichelli 2018, pp. 107-128.
- BORSELLINO P., *Libertà, giustificazione della pena e metodo delle discipline penali in Calderoni*, in *Rivista Critica di Storia della Filosofia*, 1979, 3, pp. 313-349, ripubblicato in S. SALARDI, M. SAPORITI (a cura di), *Le tecnologie 'moralì' emergenti e le sfide etico-giuridiche delle nuove soggettività*, Torino, Giappichelli, 2020, pp. 90-124.
- BORSELLINO P., *La prospettiva della filosofia del diritto*, in B. LIBERALI e L. DEL CORONA (a cura di), *Diritto e Valutazioni Scientifiche*, Torino, Giappichelli, 2022, pp. 37-53.
- BOSTROM N., *Superintelligenza. Tendenze, pericoli, strategie*, Torino, Bollati-Boringhieri, 2018.
- BUCHANAN A., *Beyond Humanity. The Ethics of Biomedical Enhancement*, Oxford, Oxford University Press, 2011.
- BUCHANAN A., *The Heart of Human Rights*, New York, Oxford University Press, 2013.
- CABITZA F., RASCHINI R., GENSINI G.F., *Unintended consequences of machine learning in medicine*, *Jama*, 2017, 318, 6, pp. 517-518.
- CALDERONI M., *I postulati della scienza positiva ed il diritto penale*, in G. PAPINI (a cura di), *Scritti di Mario Calderoni*, vol. 1, Società Anon, Editrice 'La Voce', Firenze, 1924.
- CANALE D., CIUNI R., FRIGERIO A., TUZET G. (a cura di), *Critical Thinking. Un'Introduzione*, Milano, Egea, 2021.
- CAPPIELLO B., CARULLO G., *Gli smart contracts: prospettive per l'impiego nel settore pubblico-privato*, in R. CAVALLO PERIN, D.U. GALETTA (a cura di), *Il diritto dell'amministrazione pubblica digitale*, Torino, Giappichelli, 2020, pp. 305-320.

- CARRER S., *Se l'amicus curiae è un algoritmo: il chiaccherato caso Loomis alla Corte Suprema del Wisconsin*, in *Giurisprudenza Penale Web*, 2019, 4.
- CASADEI T., PIETROPAOLI S. (a cura di), *Diritto e Tecnologie Informatiche*, Milano, Wolters Kluwer, 2021.
- CASADEI T., PIETROPAOLI S., *Intelligenza artificiale: fonte o confine del diritto?*, in ID. (a cura di), *Diritto e Tecnologie Informatiche*, Milano, Wolters Kluwer, 2021, pp. 219-232.
- CASTELLS M., *The rise of the Network Society*, Oxford, Oxford University Press, 2000.
- CATANZARITI M., *Etica "artificiale": un nuovo modello regolatorio?*, in *Ars Interpretandi*, 2021, 1, pp. 165-179.
- CERETTI A., *È possibile parlare di responsabilità morale in criminologia?*, in A. CERETTI, I. MERZAGORA (a cura di), *Criminologia e responsabilità morale*, Padova, Cedam, 1990, pp. 241-265.
- CHIRIATTI M., *Incoscienza artificiale. Come fanno le macchine a prevedere per noi*, Roma, LUISS University Press, 2021.
- CLYNES M.E., KLINE N.S., *Cyborgs in Space*, in G. HABLES GRAY, H.J. FIGUEROA-SARRIERA, S. MENTOR (eds.), *The Cyborg Handbook*, New York, Routledge, 1995, pp. 30-31.
- COHEN I.B., *La rivoluzione nella scienza*, Milano, Longanesi, 1988.
- CONRAD P., *The Medicalization of Society. On Transformation of Human Conditions into Treatable Disorders*, Baltimore, John Hopkins University Press, 2007.
- DE CARO M., MACARTHUR D. (a cura di), *Hilary Putnam. La filosofia nell'età della scienza*, Bologna, Il Mulino, 2012.
- DE MAGLIE C., *L'etica e il mercato*, Milano, Giuffrè, 2002.
- DESCARTES R., *Discorso sul metodo*, Brescia, Editrice la Scuola, 1983.
- DEWEY J., *Alcuni stadi del pensiero logico*, in R. FREGA (a cura di), *Logica Sperimentale, Teoria Naturalistica della Conoscenza e del Pensiero*, Macerata, Quodlibet, 2008, pp. 3-22.
- DEWEY J., *La teoria sperimentale della conoscenza*, in R. FREGA (a cura di), *Logica Sperimentale, Teoria Naturalistica della Conoscenza e del Pensiero*, Macerata, Quodlibet, 2008, pp. 77-95.
- DIGNUM V., *Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way*, Cham, Springer, 2019.
- DI PORTO F., *La regolazione di fronte alle sfide dell'ICT e dell'Intelligenza Artificiale*, in R. CAVALLO PERIN, D.U. GALETTA (a cura di), Torino, Giappichelli, 2020, pp. 277-303.
- DI TANO F., *Protezione dei dati personali e ricerca scientifica: un rapporto controverso ma necessario*, in *Biolaw Journal-Rivista di Biodiritto*, 2022, 1, pp. 71-98.
- DOALLO N.M., *El derecho al consentimiento informado del paciente. Una perspectiva ius-fundamental*, Granada, Editorial Comares, 2021.
- DREYFUS H.L., *What Computers Still Can't Do: A Critique of Artificial Reason*, Cambridge, The MIT Press (versione originale 1972), 1992.
- DUDZINSKI D.M., HÉBERT P.C., FOGLIA M.B., GALLAGHER T.H., *The disclosure dilemma. Large-scale adverse events*, in *New England Journal of Medicine*, 210, 363, 10, pp. 978-986.
- DUTTGE G., *Das Recht auf Nichtwissen in einer informationell vernetzten Gesundheitsversorgung*, in *Medizinrecht*, 2016, 34, pp. 664-669.

- EBERS M., *Regulating AI and Robotics. Ethical and Legal Challenges*, in M. EBERS, S. NAVAS (eds.), *Algorithms and Law*, Cambridge, Cambridge University Press, 2020, pp. 37-99.
- EISSA T.L., *Gesünder, intelligenter, perfekt? Selbstgestaltung durch Enhancement im Kontext pluralistischer Ethik*, Freiburg, Verlag Karl Alber, 2014.
- FACCHI A., *Breve storia dei diritti umani*, Bologna, Il Mulino, 2007.
- FAINI F., PIETROPAOLI S., *Scienze giuridica e tecnologie informatiche*, Torino, Giappichelli, 2021.
- FARALLI C. (a cura di), *Consenso informato in medicina. Aspetti etici e giuridici*, numero monografico di *Salute e Società*, 2012, 3, pp. 1-208.
- FARALLI C. (a cura di), *Forum Vulnerabilità e Nuove Tecnologie*, in *Notizie di Politeia*, 2019, 136, pp. 5-108.
- FARANO A., *Neuroscienze e diritto: un primo bilancio*, in S. SALARDI, M. SAPORITI (a cura di), *Le tecnologie 'moralì' emergenti e le sfide etico-giuridiche delle nuove soggettività*, Torino, Giappichelli, 2020, pp. 42-51.
- FERRAJOLI L., *Principia Juris. Teoria del diritto e della democrazia*, voll. I-II, Roma-Bari, Laterza, 2007.
- FERRAJOLI L., *Poteri selvaggi*, Roma-Bari, Laterza, 2011.
- FERRAJOLI L., *Dei diritti e delle garanzie. Conversazione con Mauro Barberis*, Bologna, Il Mulino, 2013.
- FIORIGLIO G., *Democrazia elettronica. Presupposti e strumenti*, Milano, Cedam-Wolters Kluver, 2017.
- FIORIGLIO G., *La società algoritmica fra opacità e spiegabilità: profili informatico-giuridici*, in *Ars Interpretandi*, 2021, 1, pp. 53-67.
- FLORIDI L., *La quarta rivoluzione. Come l'infosfera sta trasformando il mondo*, Milano, Raffaello Cortina, 2017.
- FLORIDI L., *The European Legislation on AI: A Brief Analysis of its Philosophical Approach*, in *Philosophy and Technology*, 2021, 34, pp. 215-222.
- FOUCAULT M., *Tecnologie del sé*, in L. MARTIN, H. GUTMAN, P.H. HUTTON (a cura di), Torino, Bollati-Boringhieri, 1992.
- GALLETTI M., *Il biopotenziamento morale e il bilanciamento tra libertà e bene*, in S. SALARDI, M. SAPORITI, M. VETIS ZAGANELLI (a cura di), *Diritti umani e tecnologie morali. Una prospettiva comparata tra Italia e Brasile*, Torino, Giappichelli, pp. 87-100.
- GIDDENS A., *The consequences of modernity*, Oxford, Polity, 1990, pp. 83-88.
- GITELMAN L., *Raw Data is an Oxymoron*, Cambridge, MIT Press, 2012.
- GOODMAN N., *I linguaggi dell'arte*, Milano, Il Saggiatore, (versione originale 1978), 2017.
- HADDADIN S., KNOBBE D., *Robotics and Artificial Intelligence*, in M. EBERS, S. NAVAS (eds.), *Algorithms and Law*, Cambridge, Cambridge University Press, 2020, pp. 1-36.
- HARARI Y.N., *Dataism is Our New God*, in *New Perspectives Quarterly*, 2017, 34, 2, pp. 36-43.
- HARDIN G., *The Tragedy of the Commons*, in *Science*, 1968, 162, 3859, pp. 1243-1248.
- HARRIS J., *How to be good. The Possibility of Moral Enhancement*, Oxford, Oxford University Press, 2016.

- HART H.L.A., *Il positivismo e la separazione tra diritto e morale*, in ID., *Contributi all'analisi del diritto*, Milano, Giuffrè, 1964, pp. 105-166.
- HART H.L.A., *Punishment and Responsibility. Essays in the Philosophy of Law*, Oxford, Clarendon Press, 1968.
- HART H.L.A., *Il concetto di diritto*, Torino, Einaudi, 1991.
- HART H.L.A., HONORÉ T., *Causation in the Law*, Oxford, Clarendon Press, 1985.
- HUME D., *An Inquiry concerning Human Understanding*, London, A. Millar, 1748.
- IENCA M., *Intelligenza2. Per un'unione di intelligenza naturale e artificiale*, Torino, Rosenberg & Sellier, 2019.
- IENCA M., *Common Human Rights Challenges Raised by Different Applications of Neurotechnologies in the Biomedical Fields*, Report Commissioned by the Committee on Bioethics (DH-BIO) of the Council of Europe, 2021.
- IENCA M., *On Neurorights*, in *Front Hum Neurosci*, 2021, 15, pp. 1-11.
- IENCA M., ANDORNO R., *Towards new human rights in the age of neuroscience and neurotechnology*, in *Life Sciences, Society and Policy*, 2017, 13, 1, pp. 1-28.
- ISTACE T., *Neurorights: the debate about new legal safeguards to protect the mind*, in *Issues in law and medicine*, 2022, 37, 1, pp. 95-114.
- JASANOFF S., *Designs on Nature*, Princeton and Oxford, Princeton University Press, 2005.
- JASANOFF S., HURLBUT G.B., SAHA K., *CRISPR Democracy. Gene Editing and the Need for Inclusive Deliberation*, in *Issues in Science and Technology*, 2015, 32, 1, pp. 25-32.
- JOHNSON-LAIRD P.N., *Deduzione, Induzione, Creatività. Pensiero Umano e Pensiero Meccanico*, Bologna, Il Mulino, 1994.
- JONAS H., *Il principio di responsabilità. Un'etica per la civiltà tecnologica*, Torino, Einaudi, 1990.
- JORI M., PINTORE A., *Manuale di teoria generale del diritto*, Torino, Giappichelli, 1995.
- KAMINSKI M., *Understanding Transparency in Algorithmic Accountability*, in W. BARFIELD (eds.), *The Cambridge Handbook of the Law of Algorithms*, Cambridge, Cambridge University Press, 2021, pp. 119-138.
- KANT I., *Critica della Ragion Pura*, Lipsia, Hartnack Riga, 1781.
- KARANASIOU A.P., *On Being Trans-Human: Commercial BCIS and the Quest for Autonomy*, in W. BARFIELD (eds.), *The Cambridge Handbook of the Law of Algorithms*, Cambridge, Cambridge University Press, 2021, pp. 757-774.
- KARNOW C.E.A., *The Opinion of Machines*, in W. BARFIELD (eds.), *The Cambridge Handbook of the Law of Algorithms*, Cambridge, Cambridge University Press, 2021, pp. 16-46.
- KARPPI T., *"The Computer Said So": On the Ethics, Effectiveness, and Cultural Techniques of Predictive Policing*, in *Social Media+Society*, 2018, April-June, 4, 2, pp. 1-9.
- KASS L.R., *Life, Liberty and the Defense of Dignity. The Challenge for Bioethics*, San Francisco, Encounter Books, 2002.
- KELLY K., *L'inevitabile. Le tendenze tecnologiche che rivoluzioneranno il nostro futuro*, Milano, Il Saggiatore, 2017.
- KELSEN H., *Lineamenti di dottrina pura del diritto*, trad. it. R. Treves, Torino, Einaudi, 1952.

- KIRCHSCHLAEGER P.G., *Digital Transformation and Ethics*, Baden-Baden, Nomos, 2021.
- KLINE R.R., *Technological Determinism*, in *International Encyclopedia of the Social & Behavioural Sciences* (Second edition), Amsterdam, Elsevier, 2015, pp. 109-112.
- LANDGREBE J., SMITH B., *Why Machines Will Never Rule the World*, New York, Routledge, 2023.
- LESSIG L., *Code is Law. On Liberty in Cyberspace*, in *Harvard Magazine*, January-February 2000, <https://www.harvardmagazine.com/2000/01/code-is-law.html>.
- LESSIG L., *Code and other laws of cyberspace*, New York, Basic Books, 2009.
- LETTIERI N., *Antigone e gli Algoritmi*, Modena, Mucchi, 2020.
- LEVY N., *Neuroethics. Ethics and the Science of the Mind*, in U. SCHÜKLENK, P. SINGER (eds.), *Bioethics. An Anthology*, New Jersey, Wiley Blackwell, 2022, pp. 861-865.
- LEWONTIN R., *Il sogno del genoma umano e altre illusioni della scienza*, Roma-Napoli, Laterza, 2000.
- LLANO ALONSO F.H., *Homo excelsior. Los límites ético-jurídicos del transhumanesimo*, Valencia, Tirant lo Blanch, 2018.
- LLANO-ALONSO F.H., *L'etica dell'intelligenza artificiale nel quadro giuridico dell'Unione europea*, in *Ragion Pratica*, 2021, 2, pp. 327-347.
- LOSANO M.G., *La lunga marcia dell'informatica nelle istituzioni italiane*, in R. CAVALLO PERIN, D.U. GALETTA (a cura di), *Il diritto dell'amministrazione pubblica digitale*, Torino, Giappichelli, 2020, pp. XXIII-LVIII.
- LOWDEN R., *Risk Assessment Algorithms: the Answer to an Inequitable Bail System?*, in *North Carolina Journal of Law & Technology*, 2018, 19, 4, pp. 221-251.
- MAGLHÃES J.C., *Do Algorithms Shape Character? Considering Algorithmic Ethical Subjectivation*, in *Social Media+Society*, 2018, April-June, 4, 2, pp. 1-10.
- MARSONET M., *I rapporti tra scienza e metafisica*, in *Acta Philosophica*, 1996, 5, 2, pp. 251-268.
- MAYER-SCHONBERGER V., CUKIER K., *Big Data: A Revolution That Will Transform How We Live, Work, and Think*, Boston, Harcourt, 2013.
- MCQUILLAN D., *People's Councils for Ethical Machine Learning*, in *Social Media+Society*, 2018, 4, 2, pp. 1-18.
- MILL J.S., *System of logic ratiocinative and inductive*, London, Longmans, Green & CO, 1886.
- MORLINO L., SORICE M. (a cura di), *L'illusione della scelta. Come si manipola l'opinione pubblica in Italia*, Roma, LUISS University Press, 2021.
- NAVAS S., *Robot Machines and Civil Liability*, in M. EBERS, S. NAVAS (eds.), *Algorithms and Law*, Cambridge, Cambridge University Press, 2020, pp. 157-173.
- NOWOTNY H., *Le macchine di Dio. Gli algoritmi predittivi e l'illusione del controllo*, Roma, LUISS University Press, 2022.
- NUSSBAUM M., *Non per profitto. Perché le democrazie hanno bisogno della cultura umanistica*, Bologna, Il Mulino, 2011.
- OBERMEYER Z., POWERS B., VOGELI C., MULLAINATHAN S., *Dissecting racial bias in an algorithm used to manage the health of populations*, in *Science*, 2019, 366, 6464, pp. 1-7.

- ORWELL G., *La politica e la lingua inglese*, in S. PERRELLA (a cura di), *Nel ventre della balena*, Milano, Bompiani, 2012.
- OTT W., *Der Rechtspositivismus*, Berlino, Duncker & Humblot, 1992.
- PACKIN N.G., LEV-ARETZ Y., *Algorithmic Analysis of Social Behaviour for Profiling, Ranking, and Assessment*, in W. BARFIELD (eds.), *The Cambridge Handbook of The Law of Algorithms*, Cambridge, Cambridge University Press, 2021, pp. 632-663.
- PALAZZANI L., *Il potenziamento umano. Tecnoscienza, etica e diritto*, Torino, Giappichelli, 2015.
- PALAZZANI L., *Tecnologie dell'Informazione e Intelligenza Artificiale. Sfide Etiche al Diritto*, Roma, Studium edizioni, 2020.
- PALAZZANI L., *AI and health: ethical aspects for regulation*, in *Teoria e Critica della Regolazione Sociale*, 2021, 1, pp. 1-16.
- PASQUALE F., *Le nuove leggi della robotica. Difendere la competenza umana nell'era dell'intelligenza artificiale*, Roma, LUISS University Press, 2021.
- PEIERLS R.E., *Le leggi della natura*, Torino, Boringhieri, 1960.
- PERSSON I., SAVULESCU J., *The Perils of Cognitive Enhancement and the Urgent Imperative to Enhance the Moral Character of Humanity*, in *Journal of Applied Philosophy*, 2008, 25, 3, pp. 162-177.
- PIROSA R., *I dilemmi della tecnoregolazione: uno sguardo critico*, in T. CASADEI e S. PIETROPAOLI (a cura di), *Diritto e Tecnologie Informatiche*, Milano, Cedam-Wolters Kluwer, 2021, pp. 275-286.
- PITTS J., *Social Control: The Concept*, in D. SILLS (eds.), *International Encyclopaedia of Social Sciences*, New York, Macmillan, 1968, vol.14.
- PIZZETTI F.G., *The Robot Sophia as a "New Citizen" of Saudi Arabia: what about granting legal personhood, "citizenship" and eventually dignity to non-human entities with artificial intelligence?*, in S. SALARDI, M. SAPORITI (forum a cura di), *Le tecnologie 'moralì' emergenti e i diritti fondamentali in Europa*, in *Notizie di Politeia*, 2019, 133, pp. 63-70.
- POPPER K., *Scienza e Filosofia*, Torino, Einaudi, 1969.
- POPPER K., *Logica della scoperta scientifica*, Torino, Einaudi, (prima edizione originale 1934), 1998.
- PRAVETTONI G., TRIBERTI S., *Il Medico 4.0: come cambia la relazione di cura nell'era delle nuove tecnologie*, Milano, Edra, 2019.
- PUNZI A., *Difettività e giustizia aumentata. L'esperienza giuridica e la sfida dell'umanesimo digitale*, in *Ars Interpretandi*, 2021, 1, pp. 113-128.
- RASOULI A., SHAO J., NEIFERT J., STEINMETZ M.P., BENZEL E., MROZ T.E., *Artificial Intelligence and Robotics in Spine Surgery*, in *Global Spine Journal*, 2021, 11, 4, pp. 556-564.
- REIDENBERG J.R., *Lex Informatica: The Formulation of Information Policy Rules Through Technology*, in *Tex. L. Rev.*, 1998, 76, 3, pp. 553-593.
- RENDA A., *Moral Machines. The emerging EU policy on "Trustworthy AI"*, in W. BARFIELD (eds.), *The Cambridge Handbook of the Law of Algorithms*, Cambridge, Cambridge University Press, 2021, pp. 667-690.
- RIEDER B., *Big Data and the Paradox of Diversity*, in *Digital Culture and Society*, 2016, 2, 2, pp. 39-54.

- RIGBY M.J., *Ethical Dimensions of Using Artificial Intelligence in Health Care*, in *AMA Journal of Ethics*, 2019, 21, 2, pp. 121-124.
- RODOTÀ S., *Se la società impone la schedatura genetica*, 2005, www.privacy.it.
- RODOTÀ S., *Il diritto di avere diritti*, Roma-Bari, Laterza, 2012.
- ROSS A., *Colpa, responsabilità e pena*, Milano, Giuffrè, 1972.
- ROSSA S., *Trasparenza e accesso all'epoca dell'amministrazione digitale*, in R. CAVALLO PERIN, D.U. GALETTA (a cura di), *Il diritto dell'amministrazione pubblica digitale*, Torino, Giappichelli, 2020, pp. 247-275.
- ROUVROY A., STIEGLER B., *Il regime di verità digitale. Dalla governamentalità algoritmica a un nuovo Stato di diritto*, in *La Deleuziana-Rivista Online di Filosofia*, 2016, 3, pp. 6-28.
- RUFFOLO U. (a cura di), *Intelligenza artificiale – Il diritto, i diritti, l'etica*, Milano, Giuffrè, 2020.
- RUFFOLO U., *La "personalità elettronica"*, in ID. (a cura di), *Intelligenza artificiale – Il diritto, i diritti, l'etica*, 2020, Milano, Giuffrè, pp. 213-234.
- RUFFOLO U., AMIDEI A., *Intelligenza artificiale, human enhancement e diritti della persona*, in U. RUFFOLO (a cura di), *Intelligenza artificiale-Il diritto, i diritti, l'etica*, Milano, Giuffrè, 2020, pp. 179-211.
- SADIN E., *Critica della ragione artificiale. Una difesa dell'umanità*, Roma, LUISS University Press, 2019.
- SALARDI S., *Test genetici tra determinismo e libertà*, Torino, Giappichelli, 2010.
- SALARDI S., *Destined to be superhuman. Moral bioenhancement and its legal viability*, in *Biolaw Journal*, 2017, 3, pp. 87-101.
- SALARDI S., *The 'Project of Moral Bioenhancement' in the European legal system. Ethical controversy and legally highly questionable*, in *Rivista di Filosofia del Diritto-Journal of Philosophy of Law*, 2018, VII, 2, pp. 241-260.
- SALARDI S., *Autonomia e responsabilità alla prova di autonomous artificial systems: Profili filosofici, etici e giuridici. Il caso delle driverless cars*, in D. CERINI, A. PISANI TEDESCO (a cura di), *Smart mobility, smart cars e intelligenza artificiale: responsabilità e prospettive*, Torino, Giappichelli, 2019, pp. 117-139.
- SALARDI S., *Lo sport come diritto umano nell'era del post-umano*, Torino, Giappichelli, 2019.
- SALARDI S., *Democrazia e nuove tecnologie: scenari passati e dell'avvenire*, in Ordines, a cura di A. LO GIUDICE, T. CASADEI, *Forum Democrazia e Parlamentarismo*, 2020, 2, pp. 397-412.
- SALARDI S., *Robótica e inteligencia artificial: retos para el Derecho*, in *Derechos y Libertades*, 2020, 42, pp. 203-232.
- SALARDI S., *When the 'age of rights' meets the 'age of technology'*, in A. D'ALOIA, M. ERRIGO (a cura di), *Neuroscience and Law. Complicated Crossings and New Perspectives*, Cham, Springer, 2020, pp. 239-255.
- SALARDI S., *L'attualità del pensiero di Mario Calderoni: questioni bioetiche di frontiera come case study*, in L. FORNI, S. SALARDI, M. SAPORITI, *Scritti in onore di Patrizia Borsellino*, Torino, Giappichelli, 2021, pp. 215-232.
- SALARDI S., *La dimensione semantica delle tecnologie intelligenti*, in *Notizie di Politeia*, 2021, 143, pp. 156-161.
- SALARDI S., SAPORITI M., *Perché l'IA non deve diventare persona. Una critica all'ineluttabile 'divenire antropomorfo' delle macchine*, in ID. (a cura di), *Le tecnologie 'mora-*

- li' emergenti e le sfide etico-giuridiche delle nuove soggettività*, Torino, Giappichelli, 2020, pp. 52-74.
- SALARDI S., SAPORITI M., *Dai principi alle regole. Appunti per una riflessione sul rapporto tra bioetica e intelligenza artificiale*, in S. SALARDI, M. SAPORITI, M. VETIS ZAGANELLI (a cura di), *Diritti umani e tecnologie morali. Una prospettiva comparata tra Italia e Brasile-Direitos Humanos e tecnologias morais. Uma perspectiva comparada entre Itália e Brasil*, Torino, Giappichelli, 2022, pp. 101-117.
- SANDEL M.J., *The Case against Perfection*, Harvard, Belknap Pr, 2009.
- SANTONI DE SIO F., VAN DE HOVEN J., *Meaningful Human Control over Autonomous Systems: A Philosophical Account*, in *Frontiers in Robotic and AI*, 2018, 5, pp. 1-14.
- SAPORITI M., *Dialettiche sovrane. Uno studio sul rapporto tra sovranità e religioni*, Torino, Giappichelli, 2019.
- SAPORITI M., *Questioni di "intelligenza politica". Prospettive europee in materia di Intelligenza Artificiale e di proceduralità algoritmica*, in *Notizie di Politeia*, 2021, 143, pp. 87-103.
- SARTOR G., *L'informatica giuridica e le tecnologie dell'informazione. Corso d'informatica giuridica*, Torino, Giappichelli, 2016.
- SCANLON T., *Moral Dimensions. Permissibility, Meaning, Blame*, Cambridge, Harvard University Press, 2008.
- SCARPELLI U., *Filosofia analitica, norme e valori*, Milano, Edizioni Comunità, 1962.
- SCARPELLI U., *La definizione nel diritto*, in ID. (a cura di), *Diritto e Analisi del Linguaggio*, Milano, Edizioni Comunità, 1976, pp. 183-197.
- SCARPELLI U., *Etica, linguaggio, ragione*, in ID., *L'etica senza verità*, Bologna, Il Mulino, 1982 pp. 49-72.
- SCARPELLI U., *John Langshaw Austin e il linguaggio comune*, in ID., *L'etica senza verità*, Bologna, Il Mulino, 1982, pp. 39-47.
- SCARPELLI U., *La «grande divisione» e la filosofia della politica*, in ID., *L'etica senza verità*, Bologna, Il Mulino, 1982, pp. 115-139.
- SCARPELLI U., *Riflessioni sulla responsabilità politica. Responsabilità, libertà, visione del mondo*, in R. ORECCHIA (a cura di), *La responsabilità politica, Diritto e Tempo*, Atti del XIII Congresso nazionale, Milano, Giuffrè, 1982, pp. 41-95.
- SCARPELLI U., *Scienza, Sapere, Sapienza*, in *Rivista Internazionale di Filosofia del Diritto*, 1986, IV Serie, LXIII, pp. 245-258.
- SCARPELLI U., *Diritti positivi, diritti naturali: un'analisi semiotica*, in S. CAPRIOLI, F. TREGGIARI (a cura di), *Diritti umani e civiltà giuridica*, Perugia, Centro Studi Giuridici e Politici, 1992, pp. 31-44.
- SCARPELLI U., *Scienza del Diritto e Analisi del Linguaggio*, in U. SCARPELLI, P. DI LUCIA, M. JORI (a cura di), *Il Linguaggio del Diritto*, Milano, LED Edizioni Universitarie, 1994, pp. 87-93.
- SCHIFF D., BORENSTEIN J., *How Should Clinicians Communicate With Patients About The Roles Of Artificially Intelligent Team Members?*, in *AMA Journal of Ethics*, 2019, 21, pp. E 138-145.
- SCHÖLKOPF B., LOCATELLO F., BAUER S., KE N.R., KALCHBRENNER N., GOYAL A., BENGIO Y., *Towards Causal Representation Learning*, in *Special Issue of Proceedings of the IEEE-Advances in Machine Learning and Deep Neural Networks*, 2021, pp. 1-23.

- SCHÜKLENK U., SINGER P. (eds.), *Bioethics. An Anthology*, New Jersey, Wiley Blackwell, 2022.
- SCHWAB K., *La quarta rivoluzione industriale*, Milano, Franco Angeli, 2016.
- SELLARS W., *Empirismo e filosofia della mente*, Torino, Einaudi, (versione originale 1956), 2004.
- SEN A., *Lo sviluppo è libertà*, Milano, Mondadori, 2014.
- SEVERINO P. (a cura di), *Intelligenza artificiale. Politica, economia, diritto, tecnologia*, Roma, LUISS University Press, 2022.
- SGAIER S.K., HUANG V., CHARLES G., *The Case for Causal AI*, in *Stanford Social Innovation Review*, 2020, 18, 3, pp. 50-55.
- SIMONCINI A., *Amministrazione digitale algoritmica. Il quadro costituzionale*, in R. CAVALLO PERIN, D.U. GALETTA (a cura di), *Il diritto dell'amministrazione pubblica digitale*, Torino, Giappichelli, 2020, pp. 1-42.
- SIMONCINI A., *Forum Law and Artificial Intelligence*, in *BioLaw Journal-Rivista di Biodiritto*, 2020, 1, pp. 498-501.
- SIMONCINI A., SUWEIS S., *Il cambio di paradigma nell'intelligenza artificiale e il suo impatto sul diritto costituzionale*, in *Rivista di Filosofia del Diritto*, 2019, VII, 1, pp. 87-106.
- SOMMAGGIO P., *Neurodiritti: tra neuroscienze e neurotecnologie*, in S. SALARDI, M. SAPORITI, M. VETIS ZAGANELLI (a cura di), *Diritti umani e tecnologie morali. Una prospettiva comparata tra Italia e Brasile*, Torino, Giappichelli, 2022, pp. 159-174.
- SONGHORIAN S., *Etica e scienze cognitive*, Roma, Carocci, 2020.
- STIRLING A., *Driving Forces in Emerging Technologies: Issues of Social Justice And Democracy In Innovation Governance*, in *Emerging Technologies and Human Rights, Proceedings of the International Conference organized by the Committee on Bioethics*, Council of Europe, 2015, pp. 8-33.
- STOCCHI F., *Presupposti etico-filosofici, casi studio e bilanciamento di interessi nella regolamentazione neurotecnologica*, in *Federalismi.it*, 2021, 29, pp. 89-120.
- SUPPES P., *A Probabilistic Theory of Causality*, Amsterdam, North Holland Publishing Company, 1970.
- TADDEO M., *Trusting Digital Technologies Correctly*, in *Minds and Machines*, 2017, 27, pp. 565-568.
- TELESIO B., *De rerum natura juxta propria principia*, 1565, riproposta a cura di V. STAMPANATO *Filosofi italiani* collezione promossa dalla Società Filosofica Italiana, Genova, A.F. Formiggini, vol. II, 1913.
- TESTART J., ROUSSEAU A., *Au péril de l'humain. Les promesses suicidaires des transhumanistes*, Paris, Éditions du Seuil, 2018.
- TURING A.M., *Macchine calcolatrici e Intelligenza*, in V. SOMENZI, R. CORDESCHI (a cura di), *La filosofia degli automi. Origini dell'intelligenza artificiale*, Torino, Paolo Boringhieri, 1986, pp. 157-183.
- VANTIN S., *Il diritto antidiscriminatorio nell'era digitale*, Milano, Wolters Kluwer, 2021.
- WINNER L., *Do Artifacts have Politics?*, in *Daedalus*, 1980, 109, 1, pp. 121-136.
- WINTHROP-YOUNG G., *Cultural techniques: Preliminary remarks*, in *Theory Culture and Society*, 2013, 30, 6, pp. 3-19.

- WINTHROP-YOUNG G., *Discourse, media, cultural techniques: The complexity of Kittler*, in *MLN*, 2015, 130, 3, pp. 447-465.
- WÖLL A., *Der Golem: Kommt der erste künstliche Mensch und Roboter aus Prag?*, in M. NEKULA, J. ROGALL (hrsg.), *Deutsche und Tschechen: Geschichte-Kultur-Politik*, München, Beck, 2001, pp. 235-245.
- ZACCARIA G., *Mutazioni del diritto: innovazione tecnologica e applicazioni predittive*, in *Ars Interpretandi*, 2021, 1, pp. 29-52.
- ZAGRBELSKY V., CHENAL R., TOMASI L., *Manuale dei diritti fondamentali in Europa*, Bologna, Il Mulino, 2016.
- ZOLA I.K., *Medicine as an Institution of Social Control*, in P. CONRAD (eds.), *The Sociology of Health and Illness: Critical Perspectives*, New York, Worth, 1972, pp. 497-506.
- ZUBOFF S., *Il capitalismo della sorveglianza. Il futuro dell'umanità nell'era dei nuovi poteri*, Roma, LUISS University Press, 2019.

Articoli di giornale

- ANDERSON B., *The End of Theory: The Data Deluge Makes the Scientific Method Obsolete*, in *Wired*, 23 giugno 2008.
- BENANTI P., MAFFETTONE S., *Algoritmi e Nuovi Strumenti di Partecipazione Democratica*, in *Corriere della Sera*, 27 luglio 2021.
- BROOKS D., *The Philosophy of Data*, in *NY Times*, 4 febbraio 2013.
- CROSS K., *Sulla coscienza delle intelligenze artificiali ci stiamo facendo le domande sbagliate*, in *Wired*, 15 giugno 2022.
- D'ALESSANDRO J., *L'educazione delle macchine*, in *La Repubblica*, 18 ottobre 2017.
- D'ALESSANDRO J., *Cara AI impara dai bambini. Le previsioni di Yann LeCun, il guru di Facebook: faremo evolvere le macchine insegnando loro a pensare come noi*, in *La Repubblica Robinson*, 7 gennaio 2018.
- DAMIANO L., DEMOUCHEL P., *Se il badante è una macchina. I nuovi robot avranno delle funzioni socialmente significative*, in *La Repubblica*, 24 aprile 2019.
- DEMUTH Y., *Die unheimliche Macht der Algorithmen*, in *Beobachter*, 26 aprile 2018.
- DOTTI L., *L'evoluzione della pratica clinica, tra intelligenza artificiale e Real World Data*, in *Wired*, 23 giugno 2021.
- FLORIDI L., *L'ultima legge della robotica*, in *La Repubblica Robinson*, 12 febbraio 2017.
- GANGADHARAN S.P., *Predictive Algorithms are not Inherently Unbiased*, in *New York Times*, 19 novembre 2015.
- HEIKKILÄ M., *Europe's AI rules open door to mass use of facial recognition, critics warn*, in *Politico*, 7 giugno 2021.

- JOHNSON K., *La Mda e la trappola dell'intelligenza artificiale senziente*, in *Wired*, 18 giugno 2022.
- LERMAN R., *Be wary of robot emotions; 'Simulated love is never love'*, in *AP News*, 26 aprile 2019.
- LO STORTO G., MANCA D., *Robot? No, Intelligenza artificiale. Le leggi di Asimov non bastano più*, in *Corriere della Sera*, 31 maggio 2021.
- MANTOVANI R., *Eva il primo robot che ricambia il sorriso degli umani*, in *Gazzetta del Sud*, 4 giugno 2021.
- MARROCCO A., *Creano poco lavoro e tanti debiti. Le facoltà umanistiche inglesi chiudono corsi e dipartimenti*, in *Huffpost*, 22 maggio 2021.
- MEDETTI S., *E se l'Europa attribuisse una "personalità" elettronica ai robot?*, in *Panorama*, 24 aprile 2018.
- METZINGER T., *EU Guidelines. Ethics Washing Made in Europe*, in *Tagesspiegel*, 8 aprile 2019.
- O'CONNEL M., *"Il transumanesimo crede che si possa sconfiggere la morte con la tecnologia"*, in *Wired*, 29 settembre 2019.
- OLDANI R., *Il dialogo tra ricerca e persone crea robot sempre più sociali*, in *Il Sole 24 Ore*, 2 settembre 2021.
- PEYRON L., *Il caso del brevetto intestato a un robot e i limiti della tecnica*, in *Il Sole 24 Ore*, 2 settembre 2021.
- SANTORO E., *Intelligenza artificiale e nuove regole*, in *Corriere Salute*, 13 maggio 2021.
- SEGANTINI E., *Intelligenza artificiale. Basta retorica, dice Varone*, in *L'Economia del Corriere della Sera*, 30 ottobre 2017.
- SCORZA G., *Smart glasses, il ritorno. Ma c'è un problema di privacy*, 20 settembre 2021.
- TREMOLADA L., *Quella strana idea di tassare i robot*, in *Il Sole 24 Ore*, 21 febbraio 2017.
- VACCARI T., *Conosciamo meglio le proteine grazie all'intelligenza artificiale DeepMind: perché è una scoperta enorme*, in *Agenda Digitale*, 2 agosto 2021.
- ZACHEO E., *La realtà degli umani negli occhi di un robot*, in *Nuovo Quotidiano di Puglia*, 28 giugno 2021.

Rapporti, pareri e risoluzioni

ATOMIUM EUROPEAN INSTITUTE, *AI4PEOPLE'S ethical framework for a good AI society: opportunities, risks, principles and recommendations*, 2018.

COMITATO NAZIONALE PER LA BIOETICA, COMITATO NAZIONALE PER LA BIOSICU-

- REZZA, LE BIOTECNOLOGIE E LE SCIENZE DELLA VITA, *Intelligenza Artificiale e Medicina: Aspetti Etici*, 2020.
- COMMISSIONE EUROPEA, *Il Libro verde sull'invecchiamento demografico*, 2021.
- COMMITTEE ON BIOETHICS (DH-BIO), *Strategic Action Plan on Human Rights and Technologies in Biomedicine (2020-2025)*, 2019.
- COUNCIL OF EUROPE STUDY DGI (2019)05, *Responsibility and AI*, 2019.
- DEUTSCHER ETHIKRAT, *Big Data und Gesundheitdatensouveränität als informationelle Freiheitsgestaltung*, 2017.
- EUROPEAN COMMISSION, *The European Digital Competence Framework for Citizens*, 2016.
- EUROPEAN COMMISSION FOR THE EFFICENCY OF JUSTICE, *European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment*, 2018.
- EUROPEAN PARLIAMENT, *Digital Sovereignty for Europe*, EPBR Ideas Papers, 2020.
- EUROPEAN PARLIAMENT, *Regulating Facial Recognition in the EU*, 2021.
- HIGH-LEVEL EXPERT GROUP ON ARTIFICIAL INTELLIGENCE, *A Definition of AI: Main Capabilities and Disciplines*, 2019.
- HIGH-LEVEL EXPERT GROUP ON ARTIFICIAL INTELLIGENCE, *Ethics Guidelines for Trustworthy AI*, 2019.
- INTERNATIONAL BIOETHICS COMMITTEE (IBC), *Report of the International Bioethics Committee of UNESCO (IBC) on the Ethical Issues of Neurotechnology*, 2021
- NUFFIELD COUNCIL ON BIOETHICS, *Artificial intelligence (AI) in healthcare and research*, 2020.
- PANEL FOR THE FUTURE OF SCIENCE AND TECHNOLOGY (STOA), *Human Enhancement Study*, 2009.
- PANEL FOR THE FUTURE OF SCIENCE AND TECHNOLOGY (STOA), *Language equality in the digital age. Toward a Human Language Project*, 2017.
- PANEL FOR THE FUTURE OF SCIENCE AND TECHNOLOGY (STOA), *A governance framework for algorithmic accountability and transparency*, 2019.
- PANEL FOR THE FUTURE OF SCIENCE AND TECHNOLOGY (STOA), *Regulating disinformation with artificial intelligence*, 2019.
- PANEL FOR THE FUTURE OF SCIENCE AND TECHNOLOGY (STOA), *Understanding algorithmic decision-making: Opportunities and challenges*, 2019.
- PANEL FOR THE FUTURE OF SCIENCE AND TECHNOLOGY (STOA), *The ethics of artificial intelligence: Issues and initiatives*, 2020.
- PANEL FOR THE FUTURE OF SCIENCE AND TECHNOLOGY (STOA), *Artificial Intelligence: From ethics to policy*, 2020.
- STEERING COMMITTEE FOR HUMAN RIGHTS IN THE FIELD OF BIOMEDICINE AND HEALTH (CDBIO), *The Impact of Artificial Intelligence on the Doctor-Patient Relationship*, 2022.

THE GENERAL CONFERENCE OF THE UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION (UNESCO), *Recommendation on the ethics of artificial intelligence*, 2021.

THE SWEDISH NATIONAL COUNCIL OF MEDICAL ETHICS, *In brief-Artificial Intelligence in healthcare*, 2020.

UNESCO INTERNATIONAL BIOETHICS COMMITTEE, *Report of the IBC on Big Data and Health*, 2017.

WORLD HEALTH ORGANIZATION, *Ethics and Governance of Artificial Intelligence for Health*, 2021.

Finito di stampare nel mese di gennaio 2023
nella Stampatre s.r.l. di Torino
Via Bologna 220

Collana del Dipartimento di Giurisprudenza dell'Università di Milano-Bicocca

Per i tipi Giuffrè

1. Luigi Fumagalli, *La responsabilità degli Stati membri per la violazione del diritto comunitario*, 2000.
2. Paolo Bonetti, *Ordinamento della difesa nazionale e costituzione italiana*, 2000.
3. Marco Antonioli, *Comunicazioni della Commissione europea e atti amministrativi nazionali*, 2000.
4. Bruno Bosco, *Pianificazione e mercato nell'economia cubana contemporanea. Analisi delle innovazioni in un sistema socialista*, 2000.
5. Aldo Chiancone, *Scritti di economia, istituzioni e finanza pubblica*, 2000.
6. Chiara Valsecchi, *Oldrado da Ponte e i suoi consilia. Un'auctoritas del primo Trecento*, 2000.
7. Nicola Bassi, *Principio di legalità e poteri amministrativi impliciti*, 2001.
8. Giulio Enea Vigevani, *Stato democratico ed eleggibilità*, 2001.
9. Marco Antonioli, *Mercato e regolazione*, 2001.
10. *The protection of the environment in a context of regional economic integration. The case of the European Community, the Mercosur and the Nafta*, edited by Tullio Scovazzi, 2001.
11. Claudio Martinelli, *L'insindacabilità parlamentare. Teoria e prassi di una prerogativa costituzionale*, 2002.
12. Barbara Biscotti, *Dal pacere ai pacta conventa. Aspetti sostanziali e tutela del fenomeno pattizio dall'epoca arcaica all'Editto giuliano*, 2002.
13. *The protection of the underwater cultural heritage. Legal aspects*, edited by Guido Camarda and Tullio Scovazzi, 2002.
14. Chiara Buzzacchi, *L'abuso del processo nel diritto romano*, 2002.
15. Margherita Ramajoli, *La connessione nel processo amministrativo*, 2002.
16. Nicola Rondinone, *Storia inedita della codificazione civile*, 2003.
17. Ettore Scimemi, *La vendita del voto. Nelle società per azioni*, 2003.
18. *La protezione del patrimonio culturale sottomarino nel Mare Mediterraneo*, a cura di Tullio Scovazzi, 2004.
19. Roberta Garabello, *La convenzione Unesco sulla protezione del patrimonio culturale subacqueo*, 2004.
20. Ubaldo G. Nannini, *Valori fondamentali e conflitto di doveri*, 2004.
21. Gabriella Citroni, *L'orrore rivelato. L'esperienza della Commissione della Verità e Ri-conciliazione in Perù: 1980-2000*, 2004.
22. *Dal Tribunale per la ex-Iugoslavia alla Corte penale internazionale*, a cura di Gianmaria Calvetti e Tullio Scovazzi, 2004.
23. Nicola Bassi, *Gli accordi fra soggetti pubblici nel diritto europeo*, 2004.
24. Matteo Gatti, *OPA e struttura del mercato del controllo societario*, 2004.
25. *La sicurezza negli aeroporti. Problematiche giuridiche ed interdisciplinari. Atti del Convegno - Milano, 22 aprile 2004*, a cura di Guido Camarda, Marco Cottone, Monica Migliarotti, 2005.
26. Irini Papanicolopulu, *Il confine marino: unità o pluralità?*, 2005.
27. *Pena, controllo sociale e modernità nel pensiero di David Garland. Atti del Convegno in onore di David Garland - Università di Milano-Bicocca, 1° marzo 2004*, a cura di Adolfo Ceretti, 2005.
28. Barbara Moretti, *La violenza sessuale tra conoscenti. Analisi giuridica e criminologica di un fenomeno poco indagato*, 2005.
29. *Quale diritto nei conflitti armati? Relazioni e documenti del ciclo di conferenze tenute nell'Università di Milano-Bicocca (marzo-maggio 2005)*, a cura di Irini Papanicolopulu e Tullio Scovazzi, 2006.

30. Giovanni Iorio, *Struttura e funzioni delle clausole di garanzia nella vendita di partecipazioni sociali*, 2006.
31. Natascia Marchei, "Sentimento religioso" e bene giuridico. *Tra giurisprudenza costituzionale e novella legislativa*, 2006.
32. *Il dominio di Venezia sul mare Adriatico nelle opere di Paolo Sarpi e Giulio Pace*, a cura di Guido Acquaviva e Tullio Scovazzi, 2007.
33. *Laicità e stato di diritto*. Atti del IV Convegno di Facoltà - Università di Milano-Bicocca, 9-10 febbraio 2006, a cura di Adolfo Ceretti e Loredana Garlati, 2007.
34. Carmela Leone, *Il principio di continuità dell'azione amministrativa. Tra operatività dell'organo, inesauribilità del potere e stabilità degli effetti*, 2007.
35. *Ordinamento penale e fonti non statali. L'impatto dei vincoli internazionali, degli obblighi comunitari e delle leggi regionali sul legislatore e sul giudice penale*. Atti delle sessioni di studio tenutesi a Milano il 21 novembre 2005, il 10 marzo e il 24 marzo 2006, a cura di Carlo Ruga Riva, 2007.
36. *Conflitti armati e situazioni di emergenza: la risposta del diritto internazionale*. Relazioni al ciclo di conferenze tenuto nell'Università di Milano-Bicocca (marzo-aprile 2006), a cura di Irini Papanicolopulu e Tullio Scovazzi, 2007.
37. *Il tribunale per la ex-Iugoslavia: l'attività svolta e il suo prossimo scioglimento*, a cura di Gianmaria Calvetti e Tullio Scovazzi, 2007.
38. Giovanni Iorio, *Le clausole attributive dello ius variandi*, 2008.
39. Claudio Martinelli, *Le immunità costituzionali nell'ordinamento italiano e nel diritto comparato. Recenti sviluppi e nuove prospettive*, 2008.
40. Loredana Garlati, *Il volto umano della giustizia. Omicidio e uccisione nella giurisprudenza del tribunale di Brescia (1831-1851)*, 2008.
41. *Immunità costituzionali e crimini internazionali*. Atti del Convegno - Milano, 8 e 9 febbraio 2007, a cura di Aldo Bardusco, Marta Cartabia, Micaela Frulli e Giulio Enea Vigevani, 2008.
42. Marco Antonioli, *Società a partecipazione pubblica e giurisdizione contabile*, 2008.
43. Mariangela Ferrari, *La compensatio luctri cum damno come utile strumento di equa riparazione del danno*, 2008.
44. Nicola Bassi, *Mutuo riconoscimento e tutela giurisdizionale. La circolazione degli effetti del provvedimento amministrativo straniero fra diritto europeo e protezione degli interessi del terzo*, 2008.
45. Delfina Boni, *Accordi OMC norme comunitarie e tutela giurisdizionale*, 2008.
46. Roberto Cornelli, *Paura e ordine nella modernità*, 2008.
47. Edoardo Dieni, *Diritto & religione vs. «nuovi» paradigmi. Sondaggi per una teoria postclassica del diritto ecclesiastico civile*, a cura di Alessandro Albisetti, Giuseppe Casuscelli, Natascia Marchei, 2008.
48. Maria Cristina Vanz, *La circolazione della prova nei processi civili*, 2008.
49. *Atti del V incontro di studio tra giovani cultori delle materie internazionalistiche*, a cura di Irini Papanicolopulu, 2008.
50. *La sécurité collective entre légalité et défis à la légalité, sous la direction de Maurizio Arcari et Louis Balmond*, 2008.
51. *Saggi in ricordo di Aristide Tanzi*, 2009.
52. Elena Marinucci, *L'impugnazione del lodo arbitrale dopo la riforma. Motivi ed esito*, 2009.
53. Sabrina Urbinati, *Les mécanismes de contrôle et de suivi des conventions internationales de protection de l'environnement*, 2009.
54. Alessandro Albisetti, *Tra diritto ecclesiastico e canonico*, 2009.
55. *Il diritto di fronte all'infamia nel diritto. A 70 anni dalle leggi razziali*, a cura di Loredana Garlati e Tiziana Vettor, 2009.
56. *La giustizia contrattuale. Itinerari della giurisprudenza italiana tra Otto e Novecento*, a cura di Giovanni Chiodi, 2009.
57. Tomaso Greco, *Le violenze psicologiche nel mondo del lavoro. Un'analisi sociologico-giuridica del fenomeno mobbing*, 2009.

58. *Le ragioni dell'uguaglianza*. Atti del VI Convegno della Facoltà di Giurisprudenza - Università degli Studi di Milano-Bicocca, 15-16 maggio 2008, a cura di Marta Cartabia e Tiziana Vettor, 2009.
59. *I diritti umani di fronte al giudice internazionale*. Atti della giornata di studio in memoria di Carlo Russo, a cura di Tullio Scovazzi, Irini Papanicopolulu e Sabrina Urbinati, 2009.
60. Giovanni Iorio, *Le trasformazioni eterogenee e le fondazioni*, 2010.
61. Lorena Forni, *La laicità nel pensiero dei giuristi italiani: tra tradizione e innovazione*, 2010.
62. *L'inconscio inquisitorio. L'eredità del Codice Rocco nella cultura processual-penalistica italiana*, a cura di Loredana Garlati, 2010.
63. Guido Camarda, Salvatore Corrieri, Tullio Scovazzi, *La formazione del diritto marittimo nella prospettiva storica*, 2010.
64. Matteo Fornari, *Il regime giuridico degli stretti utilizzati per la navigazione internazionale*, 2010.
65. Oliviero Mazza, *Il garantismo al tempo del giusto processo*, 2011.
66. *La condizione giuridica di Rom e Sinti in Italia*, 2 Tomi. Atti del Convegno internazionale - Università degli Studi di Milano-Bicocca, 16-18 giugno 2010, a cura di Paolo Bonetti, Alessandro Simoni e Tommaso Vitale, 2011.
67. Aldo Cenderelli, *Scritti romanistici*, a cura di Chiara Buzzacchi, 2011.
68. *Questions de droit international autour de l'Avis consultatif de la Cour internationale de Justice sur le Kosovo*. International Law Issues Arising from the International Court of Justice Advisory Opinion on Kosovo, sous la direction de/edited by Maurizio Arcari et Louis Balmond, 2011.
69. Andrea Massironi, *Nell'officina dell'interprete. La qualificazione del contratto nel diritto comune (secoli XIV-XVI)*, 2012.
70. *La responsabilità sociale d'impresa in tema di diritti umani e protezione dell'ambiente. Il caso dell'India*, a cura di Maria-rosa Cutillo, Rebecca Faugno e Tullio Scovazzi, 2012.
71. *Il patrimonio culturale intangibile nelle sue diverse dimensioni*, a cura di Tullio Scovazzi, Benedetta Ubertazzi e Lauso Zagato, 2012.
72. Alfredo Marra, *Il termine di decadenza nel processo amministrativo*, 2012.
73. *Diritto internazionale e bombardamenti aerei*, a cura di Massimo Annati e Tullio Scovazzi, 2012.
74. *La fabbrica delle interpretazioni*, a cura di Barbara Biscotti, Patrizia Borsellino, Valerio Pocar e Domenico Pulitanò, 2012.
75. Carmela Leone, *Gli impegni nei procedimenti antitrust*, 2012.
76. Valentina Piccinini, *I debiti di valore*, 2012.
77. Carlo Ruga Riva, *Diritto penale, regioni e territorio. Tecniche, funzione e limiti*, 2012.
78. Paolo Rondini, *Ex sententia animi tui. La prova legale negativa nell'età della codificazione*, 2012.
79. Giovanni Iorio, *Ritardo nell'adempimento e risoluzione del contratto*, 2012.
80. Mariagrazia Rizzi, *Imperator cognoscens decrevit. Profili e contenuti dell'attività giudiziaria imperiale in età classica*, 2012.
81. Giandomenico Dodaro, *Uguaglianza e diritto penale. Uno studio sulla giurisprudenza costituzionale*, 2012.
82. *Friedrich Carl von Savigny. Le questioni di principio concernenti un nuovo regolamento del processo penale*, edizione e traduzione italiana a cura di Paolo Rondini, 2012.
83. *Processo e informazione*, a cura di Loredana Garlati e Giulio Enea Vigevani, 2012.
84. Diletta Tega, *I diritti in crisi. Tra Corti nazionali e Corte europea di Strasburgo*, 2012.
85. Alessandra Donati, *Law and art: diritto civile e arte contemporanea*, 2012.
86. Diana Cerini, *Sovraindebitamento e consumer bankruptcy: tra punizione e perdono*, 2012.
87. Federico Furlan, *Presidente della Repubblica e politiche di sicurezza internazionale tra diarchia e garanzia*, edizione emendata, 2013.

88. Alessandra Daccò, «Diritti particolari» e recesso dalla s.r.l., 2013.
89. *Il ruolo del giudice nel rapporto tra i poteri*, a cura di Giovanni Chiodi e Domenico Pulitanò, 2013.
90. Alessandro Squazzoni, *Declinatoria di giurisdizione ed effetto conservativo del termine*, 2013.
91. Guido Acquaviva, *La repressione dei crimini di guerra nel diritto internazionale e nel diritto italiano*, 2014.
92. Michele Saporiti, *La coscienza disubbidiente. Ragioni, tutele e limiti dell'obiezione di coscienza*, 2014.
93. Ilario Alvino, *Il lavoro nelle reti di imprese: profili giuridici*, 2014.
94. *Il lavoro nei trasporti. Profili giuridici*, a cura di Marco Cottone, 2014.
95. *La responsabilità sociale d'impresa in tema di diritti umani e protezione dell'ambiente: il caso del Perù*, a cura di Mariarosa Cutillo, Fabián Novak e Tullio Scovazzi, 2014.
96. *La restituzione dei beni culturali rimossi con particolare riguardo alla pratica italiana*, a cura di Tullio Scovazzi, 2014.

Per i tipi Giappichelli

- 96^{bis}. *L'eredità di Uberto Scarpelli*, a cura di Patrizia Borsellino, Silvia Salardi, Michele Saporiti, 2014.
97. Silvia Salardi, *Discriminazioni, linguaggio e diritto. Profili teorico-giuridici. Dall'immigrazione agli sviluppi della tecno-scienza: uno sguardo al diritto e al suo ruolo nella società moderna*, 2015.
98. Simone Varva, *Il licenziamento economico. Pretese del legislatore e tecnica del giudizio*, 2015.
99. Lorenzo Natali, *Green criminology. Prospettive emergenti sui crimini ambientali*, 2015.
100. *Diritto e processo: rapporti e interferenze*, a cura di Filippo Danovi, 2015.
101. *La richiesta di pareri consultivi alla Corte di Strasburgo da parte delle più alte giurisdizioni nazionali. Prime riflessioni in vista della ratifica del Protocollo 16 alla Convenzione europea dei diritti dell'uomo*, a cura di Elisabetta Lamarque, 2015.
102. Nicola Ferri, *Conflicts over the conservation of marine living resources: third states, governance, fragmentation and other recurring issues in international law*, 2015.
103. Elena Depetris, *La responsabilità civile delle agenzie di rating del credito nella disciplina italiana ed europea*, 2015.
104. *Cibo e acqua. Sfide per il diritto contemporaneo. Verso e oltre Expo 2015*, a cura di Barbara Biscotti e Elisabetta Lamarque, 2015.
105. *Dialogando con Beccaria. Le stagioni del processo penale italiano*, a cura di Giovanni Chiodi e Loredana Garlati, 2015.
106. Federico Ferraris, «Rationing Justice». *La selezione dei ricorsi nelle Corti supreme di Stati Uniti e Italia*, 2015.
107. *L'avvocato nella società contemporanea*, a cura di Diana Cerini e Franco Scarpelli, 2015.
108. *Norberto Bobbio: rigore intellettuale e impegno civile*, a cura di Michele Saporiti, 2016.
109. *Princeps legibus solutus*, a cura di Alberto Maffi, 2016.
110. Lorena Forni, *La sfida della giustizia in sanità. Salute, equità, risorse*, 2016.
111. *L'immigrazione irregolare via mare nella giurisprudenza italiana e nell'esperienza europea*, a cura di Amedeo Antonucci, Irini Papanicolopulu, Tullio Scovazzi, 2016.
112. Serena Crespi, *La mobilità delle società nell'Unione europea*, 2016.
113. *La Scozia nella costituzione britannica. Storia, idee, devolution in una prospettiva comparata*, a cura di Claudio Martinelli, 2016.
114. Stefano Pellegatta, *Unità e pluralità nella obbligazione solidale passiva*, 2016.
115. *L'uso della lingua negli atti e nella comunicazione dei poteri pubblici italiani*, a cura di Paolo Bonetti, 2016.
116. Nicola Canzian, *La reviviscenza delle norme nella crisi della certezza del diritto*, 2017.
117. Massimiliano Dova, *Pena prescrittiva e condotta reintegratoria*, 2017.
118. Paolo Zicchittu, *Le "zone franche" del potere legislativo*, 2017.
119. *La legge sugli ecoreati due anni dopo. Un*

- dialogo tra dottrina e giurisprudenza, a cura di Carlo Ruga Riva, 2017.
120. Alfredo Marra, *L'amministrazione imparziale*, 2018.
 121. *Sicurezza e libertà in tempi di terrorismo globale*, a cura di Lorena Forni e Tiziana Vettor, 2017.
 122. Martino Zulberti, *Le impugnazioni delle decisioni equitative*, 2018.
 123. Antonio Angelucci, *Dietro la circoscrizione. La sfida della cittadinanza e lo spazio di libertà religiosa in Europa*, 2018.
 124. Giulio Enea Vigevani, *I media di servizio pubblico nell'età della rete. Verso un nuovo fondamento costituzionale, tra autonomia e pluralismo*, 2018.
 125. Bruno Inzitari, *Banca, crisi e responsabilità. Scritti scelti dal 1973 al 2016*, 2018.
 126. Simone Gianello, *Funzioni e responsabilità del Capo dello Stato nelle giurisprudenze costituzionali*, 2018.
 127. Alberto Villa, *Overruling processuale e tutela delle parti*, 2018.
 128. Michele Saporiti, *Dialettiche sovrane. Uno studio sul rapporto tra sovranità e religioni*, 2019.
 129. *Kállistos Nòmos. Scritti in onore di Alberto Maffi*, a cura di Barbara Biscotti, 2018.
 130. Monica Delsignore, *L'amministrazione ricorrente. Considerazioni in tema di legittimazione nel giudizio amministrativo*, 2020.
 131. *Political and legal aspect of Italian colonialism in Somalia*, edited by Elena Carpanelli and Tullio Scovazzi, 2020.
 132. *Smart mobility, smart cars e intelligenza artificiale: responsabilità e prospettive*, a cura di Diana Cerini e Andrea Pisani Tedesco, 2019.
 133. Palmina Tanzarella, *Discriminare parlando. Il pluralismo democratico messo alla prova dai discorsi d'odio razziale*, 2020.
 134. *Città, cittadini, conflitti. Il diritto alla prova della dimensione urbana*, a cura di Barbara Biscotti, Roberto Cornelli, Paolo Rondini, Carlo Ruga Riva, Alessandro Squazzoni, 2020.
 135. Ilaria Tani, *Le baie storiche. Un'anomalia nel rapporto tra terra e mare*, 2020.
 136. *Il diritto delle università nella giurisprudenza a dieci anni dalla legge n. 240/2010*, a cura di Alfredo Marra, 2020.
 137. *Scritti in onore di Patrizia Borsellino*, a cura di Lorena Forni, Silvia Salardi e Michele Saporiti, 2021.
 138. Natascia Marchei, *Il giudice civile e la nullità del matrimonio canonico trascritto*, 2021.
 139. *Lo statuto dei partiti politici tra diritto pubblico e diritto privato*, a cura di Giovanni Iorio e Paolo Zicchittu, 2021.
 140. Federico Furlan, *Il Presidente della Regione 2.0 (tra Costituzione, fonti regionali e prassi)*, 2021.
 141. *Università e anticorruzione*, a cura di Alfredo Marra e Margherita Ramajoli, 2022.
 142. Andrea Pisani Tedesco, *Il problema della responsabilità civile compensativa. Studio per un rimedio risarcitorio effettivo*, 2022.
 143. Francesco Tesauro. *Scritti scelti di diritto tributario. Vol. I. Principi e regole. Vol. II. Il processo*, a cura di Franco Fichera, Maria Cecilia Fregni, Nicola Sartori, 2022.
 144. Luca Galli, *La coprogrammazione e la progettazione dei servizi di integrazione dei migranti. Paradigmi di coinvolgimento della società civile nei percorsi di inclusione sociale*, 2022.
 145. Alessandro Semprini, *I rapporti fiduciari con funzione successoria. Contributo allo studio dei negozi con efficacia post mortem*, 2022.
 146. *Riflessioni sulla giustizia penale. Studi in onore di Domenico Pulitanò*, 2022.
 147. Silvia Salardi, *Intelligenza artificiale e semantica del cambiamento: una lettura critica*, 2023.
 148. Martinelli, *La Brexit e la Costituzione britannica. Come e perché il Regno Unito è uscito dall'Unione Europea*, 2023.
 149. Alberto Villa, *La graduazione delle questioni di merito. Ammissibilità e profili dinamici*, 2023.

