

Ricerche giuridiche

*nuovissima serie*

**Ricerche giuridiche**  
**Collana diretta da**  
A. CELOTTO, F. LIGUORI, L. ZOPPOLI

**Comitato Scientifico**

D. Amoroso, I. Caracciolo, M. Delfino, L. Fernandez Del Moral Dominguez  
M. Di Francesco Torregrossa, M. D'Onghia  
F. Galgano, L. Gatt, A. Guardiano, M. Iovane, V. Luciani  
R. Mastroianni, G. Montedoro, M. Orlandi, A. Papa  
A. Patroni Griffi, S. Prisco, R. Spagnuolo Vigorita, A. Zito

**Massimo Farina**

**MACCHINE CHE CREANO,  
DIRITTI CHE VACILLANO**

*Ontologia giuridica della creatività artificiale*

EDITORIALE SCIENTIFICA

Il volume è stato oggetto di procedura di doppio referaggio cieco  
(double blind peer review).

Volume finanziato nell'ambito del Progetto di ricerca START-UP - D.M.  
737/2021 – “HOLMES” (Harmonizing Ownership and Legal Measures for Ethical AI  
Systems) - ANNUALITÀ 2023 (CUP: F25F21002720001), con responsabile scientifico  
Prof. Massimo Farina



*Pubblicazione con licenza Open Access GOLD*

*Proprietà letteraria riservata*

© Copyright 2026 Editoriale Scientifica s.r.l.  
Via San Biagio dei Librai, 39 – 80138 Napoli  
[www.editorialescientifica.it](http://www.editorialescientifica.it) [info@editorialescientifica.com](mailto:info@editorialescientifica.com)

ISBN 979-12-235-0517-5

## INDICE

INTRODUZIONE	7
--------------	---

### PARTE PRIMA - *Fondamenti teoretici*

#### CAPITOLO I

##### L'IA COME QUESTIONE FILOSOFICO-GIURIDICA

1. Premessa	17
2. Metodologia della ricerca	21
3. La creatività come concetto giuridicamente rilevante	28
4. Il problema ontologico dell'intelligenza non umana nell'ordinamento giuridico	39

#### CAPITOLO II

##### ONTOLOGIA GIURIDICA DEL DIRITTO D'AUTORE

1. Origini gius-filosofiche della proprietà intellettuale. Riprendendo le teorie di Locke, Hegel e Bentham	43
2. L'autore come soggetto giuridico privilegiato	53
3. I confini filosofici tra opera protetta e dominio pubblico	61

### PARTE SECONDA - *IA e crisi del paradigma giuridico*

#### CAPITOLO III

##### LA SFIDA ONTOLOGICA DELL'IA GENERATIVA AL DIRITTO D'AUTORE

1. L'IA generativa come nuovo attore del processo creativo	73
2. La crisi del concetto di originalità nell'era algoritmica	82
3. Il problema dell'intenzionalità nelle creazioni generate dall'IA	83
4. La dimensione temporale della creatività artificiale	90

#### CAPITOLO IV

##### SOGGETTIVITÀ GIURIDICA E IA

1. Teorie della personalità giuridica applicabili all'IA	99
2. L'attribuzione dei diritti nell'ecosistema dell'IA generativa	104
3. Diritti fondamentali e IA	106

PARTE TERZA - *Dimensioni etiche e scenari giuridici*

## CAPITOLO V

## ETICA DELL'IA E ORDINAMENTO GIURIDICO

- |   |     |
|---|-----|
| 1. Fondamenti etici della regolamentazione dell'IA                  | 113 |
| 2. La responsabilità etica nella creazione e nell'uso di sistemi IA | 120 |
| 3. Questioni etiche specifiche nell'utilizzo creativo dell'IA       | 124 |
| 4. Verso un modello di ricerca etica e responsabile...              | 128 |

## CAPITOLO VI

## GIUSTIZIA DISTRIBUTIVA E IA GENERATIVA

- |   |     |
|---|-----|
| 1. Teorie della giustizia applicate all'economia della conoscenza | 131 |
| 2. Concentrazione del potere e asimmetrie nell'economia dell'IA   | 137 |
| 3. Diritti di remunerazione e modelli compensativi                | 140 |

PARTE QUARTA - *Prospettive regolative e traiettorie di ricerca*

## CAPITOLO VII

## MODELLI NORMATIVI PER L'IA GENERATIVA

- |  |     |
|--|-----|
| 1. Approcci regolativi comparati   | 145 |
| 2. Principi filosofico-giuridici per una regolamentazione dell'IA creativa | 155 |
| 3. Strumenti normativi innovativi  | 158 |

## CAPITOLO VIII

## SCENARI FUTURI E PROSPETTIVE DI RICERCA

- |   |     |
|---|-----|
| 1. Evoluzioni tecnologiche prevedibili e loro impatto giuridico   | 163 |
| 2. Ripensare i fondamenti del diritto d'autore                    | 168 |
| 3. L'intelligenza umana e quella artificiale: confini e relazioni | 172 |

*Conclusioni* 177

*Bibliografia* 183

*Fonti normative e giurisprudenziali* 191

*Indice analitico* 193

*Glossario dei termini tecnici* 197

## INTRODUZIONE

Quando, nell'ottobre del 2022, *DALL-E 2*, *Midjourney* e *Stable Diffusion* sono diventati accessibili al grande pubblico pochi avrebbero immaginato di trovarsi di fronte ad una invenzione (e, in senso proprio, a una vera e propria rivoluzione) paragonabile – per impatto e portata culturale – a quella registratasi con l'avvento della stampa oppure della fotografia. Al di là delle attese e anche delle percezioni, resta il fatto – che grazie a quelle applicazioni e nel giro di pochi mesi – milioni di persone in tutto il mondo hanno iniziato a generare immagini di qualità straordinaria semplicemente descrivendo ciò che desideravano vedere. Simultaneamente, *ChatGPT* dimostrava capacità linguistiche tali da produrre saggi argomentativi, componimenti poetici e analisi critiche che spostano, e sfidano, i confini tradizionali tra intelligenza umana e artificiale. Ovviamente, ciò che ha reso questi eventi particolarmente significativi non è la mera novità tecnologica – dopotutto, la storia è stata costellata di strumenti<sup>1</sup> che ambiscono e/o che tentano di emulare la creatività umana – ma il fatto che, per la prima volta, l'umanità si è trovata a misurarsi (e a confrontarsi) con sistemi di IA incapaci di esibire quella che, a tutti gli effetti pratici, integra una forma autonoma di creatività. Vale a dire, ostentando una prerogativa che – almeno sino ad allora – era stata considerata quale esclusivo (e distintivo) appannaggio dell'essere umano.

Le ricadute che affiorano da questa presa d'atto fattuale sono tanto intuitive e semplici da formulare quanto foriere di implicazioni rilevanti sotto il profilo del diritto d'autore; in modo particolare, ci si chiede se una macchina possa essere considerata autrice di un'opera di ingegno. Vien da sé che se si accetta che la creatività non sia più un monopolio esclusivo dell'intelligenza umana-biologica, l'intero edificio concettuale

<sup>1</sup> Imprescindibile, qui, il rinvio alla riflessione fondamentale di WALTER BENJAMIN (*L'opera d'arte nell'epoca della sua riproducibilità tecnica. Arte e società di massa*, trad. it., Torino, Einaudi, 1991). Inoltre, a proposito dell'evoluzione del rapporto con l'opera di ingegno, cfr. A.C. AMATO MANGIAMELI, *Sul diritto d'autore e la proprietà intellettuale, ovvero l'opera dell'ingegno tra dono, diritto e tecnica*, in EAD., *Informatica giuridica. Appunti e materiali ad uso di lezioni*, Torino, Giappichelli, 2015, pp. 297 ss.

su cui da sempre si erge il diritto d'autore moderno appare minacciato nella struttura e nella sua solidità da crepe profonde. Del resto, il sistema giuridico che si è sviluppato dal Settecento in poi è interamente fondato sul presupposto antropocentrico – che oggi appare improvvisamente fragile – secondo cui solo gli esseri umani possono creare opere meritevoli di protezione. Una creatività, quindi, che è manifestazione di immaginazione e di intenzionalità (intese quali facoltà specificamente umane) e, con riferimento alla quale, il riconoscimento giuridico serve essenzialmente a tutelare quel peculiare legame personale che si dà tra il creatore e la sua creazione. Il punto è che se questo legame – almeno apparentemente – può sussistere anche tra un algoritmo e i prodotti che genera, oppure se tale legame diventa così mediato e distribuito da risultare evanescente e irriconoscibile, allora le categorie giuridiche tradizionali rischiano di vacillare e di rivelarsi inadeguate.

Le tante istanze e le possibili criticità non investono soltanto la questione dell'autorialità – per quanto essa sia centrale – ma si estendono in direzioni molteplici e interconnesse. Vi è, innanzitutto, il problema – soltanto apparentemente tecnico ma in realtà profondamente normativo – dell'addestramento dei modelli generativi. Questi sistemi, invero, richiedono *dataset* di dimensioni colossali, composti da miliardi di immagini, testi e composizioni musicali, gran parte dei quali è protetta da *copyright*. Ora, dal canto loro, gli sviluppatori sostengono che l'apprendimento automatico da opere esistenti non costituisca una violazione dei diritti d'autore, in quanto si tratterebbe di un processo analogo a quello compiuto da un artista umano che può studiare e trarre ispirazione da opere altrui senza per questo incorrere in plagio. Diversamente, gli autori e gli editori, dal canto loro, obiettano che si tratta di un'appropriazione massiva e non autorizzata del frutto del loro lavoro creativo che si realizza su scala e con modalità radicalmente diverse dall'apprendimento umano. Al cuore di questo disaccordo si collocano due concezioni divergenti in merito a: cosa significhi effettivamente utilizzare un'opera altrui; quale sia il confine tra ispirazione legittima e sfruttamento illecito; come si debba bilanciare l'interesse pubblico-sociale verso l'implementazione e la diffusione del progresso tecnologico con la doverosa tutela dei diritti individuali.

Accanto a questa prospettiva giuridica strettamente connessa alla protezione autorale, emerge con forza crescente una questione più ampia e trasversale – di giustizia distributiva – che trascende i confini del diritto d'autore in senso stretto. Si badi: l'economia dell'IA generativa è caratterizzata da una concentrazione di potere che pochi settori tecnologici hanno conosciuto in passato, le barriere di ingresso sono notevoli. Svi-

luppate modelli all'avanguardia richiede, infatti, investimenti nell'ordine delle centinaia di milioni di dollari, accesso a capacità computazionali immense e competenze scientifiche rare. Il risultato è una situazione di autentico oligopolio<sup>2</sup> in cui poche imprese – *OpenAI, Google, Meta, Anthropic* – controllano le infrastrutture nevralgiche della creatività algoritmica. Questa concentrazione – come è intuitivo per il filosofo del diritto e per l'esperto di informatica giuridica – solleva questioni che vanno ben oltre l'efficienza economica. È necessario comprendere *chi decide* quali contenuti possono essere generati e quali no, *secondo quali criteri* vengono moderati gli *output*, e *come garantire* che i benefici dell'innovazione non si accumulino esclusivamente nelle mani di pochi attori privilegiati. Da questo punto di vista, invero, ciò che si profila all'orizzonte è il rischio concreto che i costi – in termini di dislocazione occupazionale, perdita di competenze e persino omologazione e livellamento culturale – ricadano sulla collettività, diversamente dai vantaggi che, invece, restano concentrati, coagulati nelle “mani” di pochissimi.

Ed è proprio su questo terreno che la riflessione giuridica si intreccia necessariamente con quella etica e filosofica. Le domande che l'IA generativa pone non ammettono risposte puramente tecniche o esclusivamente normative. Richiedono, invece, un ripensamento dei fondamenti stessi della creatività come la intendiamo. Diventa necessario riflettere sul significato stesso di creatività, una nozione che – nel momento in cui la novità può essere prodotta algoritmicamente ed in quantità pressoché infinite – deve per forza di cose essere risemantizzata. Il fulcro della sfida con cui i giuristi sono chiamati a misurarsi è individuato dalla duplice urgenza: da un lato, di preservare quello spazio di autenticità espressiva che da sempre consideriamo costitutivo della condizione umana, e, da un altro lato, di non abbandonarsi a un umanesimo nostalgico che rifiuta *a priori* qualsiasi forma di intelligenza non biologica. Inoltre, è importante comprendere in che modo il diritto – per definizione *socionomico*<sup>3</sup> – possa (e debba) continuare a svolgere la sua funzione ontologica di coordinamento dei diritti e delle posizioni individuali nel momento in cui le tradizionali categorie attraverso cui opera – autore, opera, creatività – vengano messe radicalmente in discussione e richiedano un ripensamento.

<sup>2</sup> Cfr. M. DURANTE, U. PAGALLO (a cura di), *La politica dei dati. Il governo delle nuove tecnologie tra diritto, economia e società*, Milano-Udine, Mimesis, 2022.

<sup>3</sup> Cfr. P. PIOVANI, *Principi di una filosofia morale*, Napoli, Morano 1972; S. COTTA, *Il diritto nell'esistenza. Linee di ontofenomenologia giuridica*, Milano, Giuffrè, 1992; ID., *Diritto, persona, mondo umano*, Torino, Giappichelli, 1989; ID., *Perché il diritto?*, Brescia, La Scuola, 1993.

Il presente lavoro monografico nasce proprio dalla consapevolezza della necessità – e dell’urgenza giur-filosofica, giuridico-informatica ed etica – di rispondere a tutte queste domande e, soprattutto, di prospettare adeguate forme di bilanciamento ed efficaci strategie di tutela. L’indagine prospettata è anche frutto della partecipazione ad una ricerca condotta nell’ambito del Progetto *Harmonizing Ownership and Legal Measures for Ethical AI Systems (HOLMES)*, elaborato in seno all’Università degli Studi di Cagliari, che è stato dedicato allo studio interdisciplinare delle implicazioni giuridiche ed etiche dell’IA. L’approccio adottato rifugge consuevolmente due tentazioni opposte – eppure in questa sede ugualmente fuorvianti – vale a dire: per un verso, quella di cedere ad un riduzionismo tecnicista che tratta l’IA come un mero problema ingegneristico risolvibile attraverso aggiustamenti normativi incrementali, senza interrogarsi sui presupposti filosofici che rendono tali aggiustamenti sensati o meno; per un altro verso, quella di eludere la tendenza ad un umanesimo conservatore che percepisce l’innovazione tecnologica essenzialmente come minaccia alla creatività autentica, senza riconoscere quelle importanti potenzialità trasformative – non necessariamente negative – che essa porta con sé.

L’idea ispiratrice è che solo un’analisi genuinamente interdisciplinare possa rendere giustizia alla complessità del fenomeno. Ciò significa integrare strumenti concettuali provenienti dalla filosofia – in particolare dalla filosofia del diritto, della mente e dall’etica normativa – con la dogmatica giuridica del diritto d’autore, senza trascurare la comprensione tecnica del modo in cui questi sistemi effettivamente funzionano e di quali sviluppi futuri siano plausibili, auspicabili e accoglibili. Non si tratta, dunque, di un mero accostamento di prospettive diverse, ma di un tentativo di costruzione teorica che prenda sul serio le tensioni e le fecondazioni reciproche tra questi diversi ambiti del sapere.

La struttura dell’opera traduce questa nuova impostazione. Nel dettaglio, nella Prima Parte si gettano le fondamenta teoretiche, che sono concettualmente imprescindibili ma che sono anche propedeutiche alle indagini successive. La trattazione trae le mosse da un capitolo – intitolato *L’intelligenza artificiale come questione filosofico-giuridica* – che inquadra l’IA come questione squisitamente filosofico-giuridica, esplorando il problema ontologico dell’intelligenza non umana e articolando la metodologia interdisciplinare adottata. L’approccio prosegue con un secondo capitolo – dedicato all’*Ontologia giuridica del diritto d’autore* – nel quale si ricostruisce la genealogia del diritto d’autore, evidenziando come questo si sia sviluppato storicamente in risposta a specifici problemi sociali ed economici e come incorpori determinati presupposti filosofici, in partico-

lare con riguardo alla teoria lockiana del lavoro, a quella hegeliana circa la personalità e, non da ultimo, alla speculazione utilitarista in ordine all'incentivo. Questa ricostruzione, ovviamente, non ha un valore meramente ricognitivo, poiché è tesa a comprendere le ragioni profonde che hanno accompagnato – e giustificato – l'emergere del diritto d'autore, rivelandosi una condizione necessaria per valutare se (e se sì e in che misura) tali ragioni conservino la loro validità originaria di fronte alle sfide dell'IA.

La Seconda Parte si incentra sulla crisi del paradigma tradizionale. Nello specifico, con il capitolo intitolato *La sfida ontologica dell'IA generativa al diritto d'autore* si riflette sulla riconfigurazione sostanziale che l'IA generativa impone al diritto d'autore. Emerge, qui, la necessità di comprendere in che senso questi sistemi possano dirsi nuovi attori del processo creativo<sup>4</sup>. Anzitutto, si rende necessario ridefinire il concetto di originalità soprattutto nel momento in cui la produzione è frutto di elaborazione statistica su *dataset* immensi. Occorre, inoltre, considerare che cosa accada al requisito dell'intenzionalità quando a guidare il processo creativo non sia una mente cosciente. Segue poi il capitolo su *Soggettività giuridica e intelligenza artificiale*, che esplora la questione della soggettività giuridica attraverso una serie di domande radicali. Qui si prende in esame la questione – decisamente nodale e controversa – del possibile riconoscimento della personalità giuridica (in dottrina, si parla di eventuale attribuzione della c.d. personalità elettronica) ad un soggetto non unitario e distribuito su molteplici attori.

La Terza Parte sposta il *focus* sulle implicazioni etico-normative. Nel capitolo dedicato a *Etica dell'intelligenza artificiale e ordinamento giuridico*, la disamina si rivolge ai fondamenti etici della regolamentazione dell'IA. Si ripercorrono, inoltre, le diverse teorie normative (dal principialismo medico-bioetico all'etica delle virtù, dal consequenzialismo utilitarista alla deontologia kantiana) prospettandone le diverse conclusioni e i vari spunti critici che possono trovare – *mutatis mutandis* – una utile riformulazione in chiave pratica. Fra le questioni emergenti e oggetto di riflessione e di dibattito: la legittimità del *training* su opere protette, la produzione di *deepfakes*, la mercificazione della creatività. Segue il capitolo *Giustizia distributiva e intelligenza artificiale generativa*, nel quale si affrontano le questioni di giustizia distributiva, mettendo in dialogo le teorie di Rawls, di Sen e di Nozick. Qui il *focus* è costituito dai problemi suscitati dalla concentrazione oligopolistica del potere tecnologico e dalla necessità di garantire forme

<sup>4</sup> Cfr. M. WEBER, *Il lavoro intellettuale come professione*, trad. it., Einaudi, Torino, 1971.

adeguate di remunerazione agli autori le cui opere contribuiscono alla creazione di valore nell'ecosistema dell'IA.

La Quarta Parte, infine, volge lo sguardo alle soluzioni possibili e agli scenari futuri. Qui troviamo il capitolo dedicato a *Modelli organizzativi per l'intelligenza artificiale generativa*, nel quale si analizzano i modelli normativi emergenti grazie a una comparazione tra i diversi approcci giurisdizionali (quello europeo che si fonda sull'*AI Act* e su una logica precauzionale; quello americano che si orienta al *fair use* e alla libertà di innovazione; quelli asiatici che presentano caratteristiche assolutamente peculiari e meritevoli di attenzione specifica). Si passa, poi, all'esame di alcuni strumenti regolativi innovativi come la *soft law*, i *regulatory sandboxes* e la standardizzazione tecnica. Non da ultimo, si illustrano delle possibili proposte (alcune delle quali elaborate nell'ambito del *Progetto HOLMES*), che includono: *guidelines* per l'utilizzo etico dei sistemi IA, *best practices* per la protezione della proprietà intellettuale, raccomandazioni legislative *de iure condendo*. L'indagine si conclude con un capitolo intitolato *Scenari futuri e prospettive di ricerca*, in cui si tratteggiano i possibili orizzonti che ci attendono: dall'IA generale a quella quantistica, dalle interfacce cervello-computer all'ibridazione umano-artificiale. Sviluppi, questi, che potrebbero richiedere ripensamenti ancora più pervasivi e radicali di quelli sollecitati con riferimento ai fondamenti del diritto d'autore<sup>5</sup>.

Nel dare alle stampe questo volume appare quanto mai necessario rivolgere un'avvertenza al lettore: l'IA rappresenta un settore in continua e tumultuosa evoluzione<sup>6</sup>, che, come tale, è suscettibile di cambiamenti e di riconfigurazioni talmente rapide da rendere qualsiasi analisi potenzialmente già desueta. Proprio a partire da questa consapevolezza, si è scelto di valorizzare la trattazione delle questioni filosofiche di fondo, piuttosto che limitarsi ad una disamina dei soli elementi tecnici contingenti. L'auspicio è che – in questo modo – i risultati prospettati possano conservare una certa rilevanza anche nel momento in cui le architetture algoritmiche prese in esame saranno superate da altre tecnologie più performanti.

Al di là dei rivolgimenti contingenti, infatti, le istanze radicali – circa la natura della creatività, in ordine al merito del riconoscimento autoriale, a proposito del bilanciamento legittimo tra innovazione tecnologica e tutela dei diritti – persisteranno comunque, indipendentemente dalle declinazioni e dalle future manifestazioni dell'IA. Del resto, in ossequio a quello che è l'approccio tipico della filosofia del diritto, la presente ricerca non ambisce (e

<sup>5</sup> N. LEVAL, "Toward a Fair Use Standard", in *Harvard Law Review*, 103 (1990), p. 1105 ss.; *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569 (1994).

<sup>6</sup> Cfr. S. COTTA, *La sfida tecnologica*, il Mulino, Bologna, 1969.

non potrebbe ambire) a fornire un catalogo di risposte “definitive” a problemi ancora largamente aperti e di là da venire. L’intento è piuttosto quello di contribuire attivamente all’elaborazione di quel processo deliberativo collettivo, attraverso il quale le società democratiche decidono consapevolmente come orientare – e come disciplinare – il proprio futuro tecnologico.

Si badi: l’IA non può – e non deve – essere percepita alla stregua di un “destino ineluttabile” che ci investe dall’esterno, essa è, più correttamente, una tecnologia che noi stessi sviluppiamo e che, quindi, possiamo consapevolmente ed efficacemente orientare verso esiti più o meno desiderabili. Questo approccio più consapevole ed *engagé* richiede una comprensione sincronica non solo (e non solamente) delle immani possibilità tecnico-pratiche, ma anche (e soprattutto) delle implicazioni normative, dei valori in gioco, nonché dei *trade-off* inevitabili. Invero, solo la ricerca rigorosa può fornire questa visione prismatica e onnicomprensiva. Tuttavia, la comprensione – per quanto imprescindibile – non è sufficiente. Occorre, infatti, anche volontà politica e la partecipazione democratica, giacché le scelte decisive non spettano (e non possono spettare in via esclusiva) agli esperti, ma alla società nel suo insieme.

Se questo volume riuscirà a contribuire – anche solo parzialmente – al dibattito pubblico, se concorrerà a rendere più consapevole la valutazione delle decisioni che ci attendono, se aiuterà ad immaginare i modi in cui l’IA potrà essere utilizzata per favorire la fioritura dell’umano (anziché la deriva verso la sua minaccia e la sua sostituzione) esso avrà adempiuto allo scopo che si è proposto il suo autore. La posta in gioco, invero, non è semplicemente la regolamentazione di una nuova tecnologia – per quanto strabiliante e potente – bensì il futuro della creatività stessa, della cultura come spazio di espressione autenticamente umana<sup>7</sup> e, forse, persino della stessa condizione umana<sup>8</sup>. In un’epoca in cui i confini tra intelligenza biologica e artificiale si fanno sempre più sfumati, porosi e labili, è quanto mai urgente e necessario promuovere una disamina critica capace di risemantizzare paradigmi e categorie tradizionali – e, laddove necessario, anche di elaborarne di nuovi – così da rispondere all’esigenza di riaffermare e garantire i diritti.

<sup>7</sup> Interessante, qui, ricordare la ricostruzione di G. SANTYANA, *The Life of Reason, or the Phase of Human Progress, Introduction and Reason in common sense*, The MIT Press, New York, 1905.

<sup>8</sup> Seppur in un contesto parzialmente diverso – quale è quello attuale – è imprescindibile il rimando alle fondamentali e intramontabili riflessioni di M. HEIDEGGER (*La questione della tecnica*, trad. it., GoWare, Firenze, 2017), H. JONAS (*Il principio responsabilità. Un’etica per la civiltà tecnologica*, trad. it., Einaudi, Torino, 2009), H. ARENDT (*Vita activa. La condizione umana*, trad. it., Bompiani, Milano, 2019).



PARTE PRIMA

*Fondamenti teoretici*



## CAPITOLO I

### L'IA COME QUESTIONE FILOSOFICO-GIURIDICA

SOMMARIO: 1. Premessa. – 2. Metodologia della ricerca. – 3. La creatività come concetto giuridicamente rilevante. – 4. Il problema ontologico dell'intelligenza non umana nell'ordinamento giuridico.

#### 1. *Premessa*

La questione della creatività artificiale si configura come una sfida ontologica radicale che investe l'essenza stessa dell'autorialità e interroga profondamente la funzione sociale del diritto<sup>1</sup>. Affrontare il fenomeno della produzione algoritmica senza una solida base filosofica significherebbe, infatti, condannare il discorso giuridico ad essere preda di un'oscillazione perenne tra intuizioni inarticolate e decisioni arbitrarie, minando alla base – oltre che la certezza e la coerenza sistematica dell'ordinamento – anche quello che è il senso stesso del diritto. Per ovviare a tale lacuna, l'indagine deve muovere dallo statuto ontologico degli attori sociali, poiché interrogarsi sulla possibilità che un'intelligenza artificiale assurga al rango di autore impone preliminarmente di definire quali siano le condizioni che permettono a un'entità di essere riconosciuta come tale dalla comunità di riferimento. Risulta, pertanto, impellente mappare le tradizioni filosofiche esistenti per rintracciare una bussola metodologica capace di orientare la riflessione oltre le contingenze eminentemente tecniche, per approdare ad una sintesi teorica che non si limiti a una sterile ricognizione, ma si faccia carico di una rigorosa opera di *design* istituzionale fondata sulla comprensione delle strutture profonde della realtà sociale, oltre che, ovviamente, del portato primigenio del fenomeno giuridico.

Nel tentativo di provare a inquadrare la creatività artificiale entro perimetri e paradigmi concettuali noti, il pensiero giuridico (e in particolar modo quello gius-filosofico) attinge alle teorie prospettate da tre grandi direttrici filosofiche, che tuttavia – pur offrendo strumenti concettuali di indubbio valore – denotano dei limiti strutturali nel momento in cui debbano essere applicate alla complessità dei sistemi generativi contemporanei.

La prima corrente teorica alla quale si rinvia è quella facente capo alla

<sup>1</sup> Cfr. S. COTTA, *Il diritto nell'esistenza. Linee di ontofenomenologia giuridica*, cit.

tradizione metafisica classica (che da Aristotele giunge fino alla Scolastica). Si tratta di quel filone speculativo che riflette sull'essenza immutabile dei fenomeni e che – interrogandosi sulla causa formale in ragione della quale un'azione viene (e può essere) considerata autenticamente creativa – procede alla fondamentale distinzione tra *poiesis* (ποίησις) e *praxis* (πράξις). Laddove la prima (ovvero la *poiesis*) è intesa come produzione/fabbricazione orientata all'opera esterna, mentre la seconda (ossia la *praxis*) individua l'azione che trova in sé il proprio fine<sup>2</sup>. Vien da sé che l'approccio metafisico – ancorato a una concezione della creatività come prerogativa esclusiva della libertà e della coscienza (retaggio della *creatio ex nihilo* teologica) – faticò ad accettare e ad integrare agenti non-umani che operano per ricombinazione di *pattern*, rischiando, evidentemente, di ipostatizzare definizioni antropocentriche che si rivelano troppo rigide per catturare la plasticità delle istituzioni.

Procedendo a grandi passi fra le epoche e rivolgendo lo sguardo a tutt'altra direzione speculativa, di segno diametralmente opposto è il paradigma positivista, il quale – rifuggendo in radice dalle questioni essenzialiste – si limita ad adottare una postura formalistico-descrittiva in aderenza alla quale la creatività è appiattita a una serie di criteri operazionali empiricamente verificabili. Tale approccio – in adesione alla prospettiva non-cognitivistica e, dunque, in ottemperanza a una netta separazione tra “essere” e “dover-essere”<sup>3</sup> – tende a privilegiare una postura formalistico-descrittiva, mostrando minore disponibilità verso una critica normativa sostanziale e rischiando, così, di ridurre il compito del giurista a quello di un osservatore della presenza o dell'assenza di protezione nel sistema normativo vigente, senza offrire sempre strumenti adeguati per valutare se tale esclusione sia coerente con le finalità del sistema giuridico stesso.

<sup>2</sup> In modo particolare, la *praxis* – proprio perché sprovvista di un esito esterno, che possa essere preliminarmente determinato e conosciuto – corrisponderebbe all'azione nel senso originario. Di qui, la percezione secondo la quale alla *praxis* non sarebbe riferibile un autore, dal momento che il soggetto (o i soggetti) della *praxis*, sono attori. La *praxis*, invero, non è mai interamente riconducibile alle intenzioni di chi la realizza ed è sottratta alla nozione di prevedibilità e di necessità. Cfr. H. ARENDT, *Vita activa. La condizione umana*, trad. it., Milano, Bompiani, 1964, pp. 11-23 e 69-76. Sulla distinzione tra *labor*, *work* e *action* e, più in generale, sulla differenza tra *animal laborans*, *homo faber* e *zoon politikon*, le pagine di Arendt restano un riferimento imprescindibile.

<sup>3</sup> Imprescindibile, il rinvio alla fondamentale ricostruzione teorica, acme e caposaldo della prospettiva positivista, di H. KELSEN, *La dottrina pura del diritto*, trad. it., Einaudi, Torino, 1966. Nel panorama italiano, e in commento, cfr.: N. BOBBIO, *Teoria dell'ordinamento giuridico*, Torino, Giappichelli, 1960; U. SCARPELLI, *Cos'è il positivismo giuridico*, Napoli, Edizioni Scientifiche Italiane, 1997.

Fra queste due visioni del mondo opposte si rinviene un'altra – e diversa – lettura, ascrivibile al criticismo kantiano, il cui metodo trascendentale non si interroga sull'essere dell'oggetto, ma sulle condizioni di possibilità del giudizio di creatività. Resta però il fatto che, anche in questo caso, le categorie concettuali della genialità e dell'immaginazione produttiva rimangono comunque sia pensate e modellate sulla razionalità umana, con una concettualizzazione che è difficilmente trasponibile a una macchina. Anche di qui, dunque, l'urgenza di una sintesi diversa.

Il superamento dei limiti e delle aporie dei modelli tradizionali trova una via d'uscita feconda nell'ontologia sociale<sup>4</sup>, che si avvicina al diritto come a un sofisticato sistema di fatti istituzionali ontologicamente dipendenti dall'intenzionalità collettiva. In questa prospettiva, risulta fondamentale la distinzione tra fatti bruti – ovvero esistenti indipendentemente dall'osservatore – e fatti istituzionali, la cui realtà è costituita da regole che non si limitano a disciplinare attività preesistenti (le cosiddette regole regolative) ma che creano la possibilità stessa dell'attività, traducendosi in regole costitutive. Il diritto d'autore non deve essere, dunque, inteso come una proprietà metafisica immanente all'opera, bensì come uno *status* attribuito secondo cui la norma non protegge una creatività naturale già data, ma costituisce lo statuto giuridico dell'opera dell'ingegno.

Integrando la teoria della documentalità elaborata da Maurizio Ferraris, si comprende che l'oggetto sociale emerge secondo la formula sintetica “*oggetto sociale = atto iscritto*”, in ragione del fatto che la fissazione o l'iscrizione su un supporto materiale diviene l'elemento costitutivo che permette all'atto di sedimentarsi nella realtà sociale. La teoria della documentalità non intende appiattare l'opera a un mero adempimento burocratico oppure a una registrazione meramente formale. Al contrario, essa individua nella fissazione dell'*output* algoritmico su un supporto (ossia nell'iscrizione) la base ontologica necessaria affinché l'atto di generazione computazionale possa essere socialmente percepito e, successivamente, essere investito dello *status* di opera intellettuale. Altrimenti detto, se il diritto d'autore per così dire “classico” è stato tradizionalmente ricondotto all'atto spirituale del soggetto umano, il diritto chiamato a confrontarsi con l'intelligenza artificiale muove più propriamente dall'atto iscritto, poiché è attraverso la documentalità che l'*agency* distribuita della macchina trova una forma stabile sulla quale il diritto può radicare l'attribuzione di responsabilità e

<sup>4</sup> Immediato il richiamo a J.R. SEARLE, *The Construction of Social Reality*, New York, Free Press, 1995. Mentre, nel panorama italiano, assai significativa la ricostruzione di M. FERRARIS, *Documentalità. Perché è necessario lasciar tracce*, Roma-Bari, Laterza, 2009.

di titolarità. In questo senso può dirsi che l'attribuzione di autorialità all'intelligenza artificiale si configura, allora, come una sorta di atto linguistico performativo la cui validità dipende dal rispetto di procedure e di criteri di riconoscimento che la comunità giuridica decide di adottare *ex ante* e, in modo particolare, in sede di *design* istituzionale.

Come è evidente, l'indagine sull'ontologia sociale – anche in considerazione della peculiare complessità dell'oggetto – deve necessariamente nutrirsi di contributi interdisciplinari che siano in grado di coniugare gli studi di filosofia della mente, quelli di filosofia morale e di filosofia della tecnica; il tutto entro la cornice di una narrazione di ampio respiro capace di gestire l'*agency* distribuita propria della rete degli attori.

Sotto il profilo della filosofia della mente, è ineludibile il confronto con la distinzione avanzata da Searle tra intenzionalità intrinseca e intenzionalità derivata. Nel dettaglio, l'esperimento della *Stanza Cinese*, a suo modo, mette in evidenza come la manipolazione sintattica di simboli non equivalga di per sé alla comprensione semantica, ponendo un problema serio per l'epistemologia giuridica. Invero, se l'opacità algoritmica impedisce di verificare la sussistenza di un'intenzione creativa consapevole, su quali basi il diritto può ancora fondare il requisito dell'originalità? Parallelamente, la filosofia morale si interroga circa i fondamenti del merito e della tutela, contrapponendo la visione lockiana del lavoro e quella hegeliana della personalità all'approccio utilitarista. Diversamente, la filosofia della tecnica e la *Actor-Network Theory* di Bruno Latour suggeriscono di riconsiderare la creatività non come un atto isolato di un soggetto sovrano, bensì come un'emergenza distribuita all'interno di una rete composta ed eterogenea di attori umani e non-umani. In tale ottica, l'insistenza del diritto su un autore individuale appare come un disallineamento ontologico rispetto alla realtà – decisamente più complessa – che è insita nella produzione algoritmica, rendendo necessaria una riflessione ben diversa che si incentra sulla distribuzione diffusa dell'*agency* tra programmatori, utenti e sistemi, che implica il superamento della tradizionale dicotomia tra agente-attore e strumento, fra creatore e opera-prodotto della creazione stessa.

La sintesi finale di questo percorso approda alla definizione di un'ontologia sociale critico-costruttiva, un approccio metodologico che, muovendo dall'esigenza di superare le limitazioni prospettive dei modelli isolati, guarda al diritto d'autore come a un'architettura di riconoscimento razionalmente fondata. Tale indagine si articola su tre livelli distinti, ma strettamente sinergici e interconnessi: *i*) un livello descrittivo-genealogico, volto a ricostruire la sedimentazione storica delle categorie autoria-

li; *ii*) un livello critico-normativo, finalizzato a vagliare la coerenza delle norme rispetto ai principi fondanti del sistema; *iii*) un livello costruttivo-progettuale, orientato alla riconfigurazione delle categorie giuridiche di fronte alle sfide dell'intelligenza artificiale.

L'attribuzione dello *status* di autore all'intelligenza artificiale (o al suo *output*) deve essere intesa come una deliberata scelta di *design* istituzionale che richiede giustificazioni razionali coerenti con le finalità proprie del diritto e della giustizia.

## 2. Metodologia della ricerca

La premessa ha messo in luce la complessità filosofica e giuridica che caratterizza il confronto tra diritto d'autore e intelligenza artificiale. Di fronte a tale scenario, risulta evidente che nessuna disciplina, laddove venga presa isolatamente, possa offrire risposte adeguate alle diverse istanze che emergono. A maggior ragione, il fenomeno dell'IA generativa necessita di una contaminazione e di una sincronia fruttuosa fra le prospettive filosofiche, giuridiche, tecnologiche ed etiche, ricondotte in un quadro comunque sia coerente e sistematico. L'ambizione di questo volume è proprio quella di sviluppare questo genere di approccio, avviando e incentivando un dialogo tra saperi diversi, quale preconditione necessaria e imprescindibile per elaborare soluzioni normative che siano, al contempo, teoricamente fondate e praticamente auspicabili, in quanto efficaci ed efficienti.

L'interdisciplinarietà, ovviamente, non può essere intesa come una semplice giustapposizione di contributi provenienti da ambiti differenti, né alla stregua di una superficiale contaminazione di linguaggi e di metodi eterogenei. Come ha osservato Edgar Morin nella sua riflessione sulla complessità, "l'interdisciplinarietà può significare semplicemente che discipline diverse si mettono attorno a un tavolo, come nazioni diverse si riuniscono all'ONU, senza fare altro che affermare ciascuna i propri diritti e le proprie sovranità rispetto alle incursioni del vicino"<sup>5</sup>. Per evitare questa semplificazione – e per scongiurare questo rischio – è necessario non solo definire quale ruolo spetti a ciascuna disciplina, ma stabilire anche come le diverse prospettive possano integrarsi senza snaturarsi e dissolversi e, soprattutto, mettere in luce quale sia il fulcro concettuale attorno al quale si coagula

<sup>5</sup> E. MORIN, *La Méthode 1: La Nature de la Nature*, Paris, Éditions du Seuil, 1977; ID., "Articuler les disciplines", in *Carrefour des sciences. Actes du colloque du Comité national de la recherche scientifique "Interdisciplinarité"*, Paris, CNRS, 1990.

e si snoda l'indagine. Un fulcro che, ad avviso di chi scrive, è individuato dalla tutela della persona e della dignità umana: parametro fondamentale e ineludibile per misurare la legittimità di ogni innovazione algoritmica<sup>6</sup>.

In prima battuta, è opportuno chiarire che cosa si intende per approccio filosofico-giuridico al fenomeno tecnologico, evidenziando tanto le potenzialità euristiche di tale prospettiva quanto i suoi limiti intrinseci. Inoltre, si rifletterà criticamente sui limiti e sulle potenzialità dell'analisi gius-filosofica dell'innovazione tecnologica, evitando, sia, l'ingenuo ottimismo di chi ritiene che la teoria possa risolvere tutti i problemi pratici, sia, lo scetticismo di chi nega qualsiasi utilità alla riflessione filosofica in ambito giuridico.

Si osservi: la posta in gioco di questa chiarificazione di stampo metodologico non è puramente accademica. In un'epoca in cui la velocità dell'innovazione tecnologica tende a superare la capacità delle istituzioni di comprenderla e di regolarla<sup>7</sup>, infatti, diventa quanto mai cruciale sviluppare metodologie di ricerca che permettano di anticipare i problemi piuttosto che limitarsi a reagire ad essi. Il diritto, tradizionalmente conservatore e orientato alla stabilità, rischia di trovarsi sempre in ritardo rispetto ai fenomeni che dovrebbe governare. L'approccio filosofico-giuridico, se ben condotto, può offrire strumenti concettuali che permettono di individuare le invarianti strutturali al di là delle contingenze tecniche, identificando così i nodi problematici prima che essi si manifestino in tutta la loro urgenza pratica. È in questa capacità anticipatrice e orientativa che risiede, in ultima analisi, il valore dell'indagine filosofica per il diritto.

L'espressione "approccio filosofico-giuridico" designa una modalità di indagine che si colloca all'intersezione tra la filosofia del diritto e la teoria giuridica *tout court*, e che si caratterizza per l'attenzione rivolta ai presupposti concettuali e normativi delle categorie giuridiche. A differenza della dogmatica giuridica – che assume come dati i concetti elaborati dal legislatore e dalla giurisprudenza limitandosi a sistematizzarli e interpretarli – l'approccio filosofico-giuridico sottopone tali concetti a indagine critica, interrogandosi sulla loro coerenza interna, riflettendo sulla loro giustificazione normativa e sulla loro adeguatezza rispetto ai fenomeni che pretendono di regolare. Nel contesto dell'intelligenza artificiale e del diritto d'autore, questo significa non limitarsi a chiedersi come

<sup>6</sup> Sul nesso tra dignità, diritti fondamentali e architettura della democrazia costituzionale, cfr. N. BOBBIO, *L'età dei diritti*, Torino, Einaudi, 1990; L. FERRAJOLI, *Principia iuris. Teoria del diritto e della democrazia*, Roma-Bari, Laterza, 2007; S. RODOTÀ, *Il diritto di avere diritti*, Roma-Bari, Laterza, 2012.

<sup>7</sup> Osservazioni, queste, che nel panorama della Filosofia del diritto prospettò – fra i primi – S. COTTA (*La sfida tecnologica*, cit.).

le norme vigenti si applichino alle opere generate dall'IA, ma spingersi a interrogarsi sulla tenuta stessa dei concetti fondamentali – nel caso di specie: la nozione di autore, di opera di ingegno, di creatività e di originalità – alla luce degli incessanti e radicali mutamenti tecnologici. Si tratta di una prospettiva che – per certi versi e *secundum quid* – si rifà anche a illustri precedenti speculativi: anzitutto la ricostruzione di Herbert Hart (che, nella sua critica al positivismo giuridico tradizionale, ha evidenziato l'importanza dell'analisi del linguaggio giuridico e dei concetti che strutturano il ragionamento dei giuristi, proponendo una metodologia che integra filosofia analitica e teoria del diritto<sup>8</sup>) e, ovviamente, anche quella di Ronald Dworkin (che, seppur da una prospettiva diversa, ha sostenuto che l'interpretazione giuridica presuppone sempre una teoria normativa circa i principi che giustificano il diritto vigente, teoria che non può essere elaborata senza il concorso della filosofia morale e politica)<sup>9</sup>. Va detto subito che l'approccio filosofico-giuridico al fenomeno tecnologico – così come viene inteso e formulato in questa sede – attinge a queste tradizioni pur senza identificarsi con nessuna di esse. In particolare, esso assume tre caratteristiche metodologiche fondamentali.

In primo luogo, la dimensione “genealogica”: prima di valutare se e come le categorie giuridiche tradizionali possano essere applicate ai nuovi fenomeni tecnologici, occorre ricostruire la genesi storica di tali categorie, comprendendo quali presupposti filosofici, sociali e politici ne hanno determinato la formazione. Come ha mostrato Michel Foucault nei suoi studi<sup>10</sup>, i concetti che appaiono naturali e necessari sono spesso il prodotto di contingenze storiche e svelarle può aprire spazi di riflessione critica e di trasformazione. Nel caso del diritto d'autore, la genealogia rivela che concetti quali quello di autore e di creatività sono inseparabili dalla teoria romantica del genio e dall'ideologia proprietaria della modernità borghese e questa consapevolezza è essenziale per valutare se tali nozioni possano – e debbano – essere preservate, modificate e/o vadano deposte e abbandonate.

In secondo luogo, la dimensione “analitico-concettuale”: l'approccio filosofico-giuridico si avvale degli strumenti della filosofia analitica per chiarificare i perimetri dei concetti giuridici, distinguere i diversi significati che essi possono assumere ed individuare le relazioni logiche che intercor-

<sup>8</sup> H.L.A. HART, *The Concept of Law*, Oxford, Oxford University Press, 1961.

<sup>9</sup> R. DWORKIN, *Law's Empire*, Cambridge (MA), Harvard University Press, 1986.

<sup>10</sup> M. FOUCAULT, “Nietzsche, Genealogy, History”, in D.F. BOUCHARD (ed.), *Language, Counter-Memory, Practice: Selected Essays and Interviews*, Ithaca (NY), Cornell University Press, 1977, pp. 139-164.

rono tra di essi<sup>11</sup>. Come insegna la tradizione analitica da Frege a Quine, molti problemi filosofici nascono da confusioni concettuali oppure da ambiguità semantiche e la loro soluzione richiede, innanzitutto, un'opera di chiarificazione linguistica. Nel contesto dell'IA e del diritto d'autore, l'analisi concettuale permette, ad esempio, di distinguere tra diverse accezioni di creatività (psicologica, storica, giuridica), di precisare cosa si intende per originalità (novità soggettiva, novità oggettiva, impronta personale) e di identificare le condizioni necessarie e sufficienti affinché un artefatto possa essere considerato opera nel senso giuridicamente rilevante.

In terzo luogo, la dimensione "normativa": l'approccio filosofico-giuridico non si limita a descrivere le categorie giuridiche esistenti o a ricostruirne la genesi, ma si interroga sulla loro giustificazione e propone criteri propedeutici alla loro revisione così come alla loro riforma e al loro eventuale superamento. A differenza dell'analisi economica del diritto, che valuta le norme principalmente in base alla loro efficienza, l'approccio filosofico-giuridico considera centrale la questione della giustizia e dei valori che il diritto è chiamato a tutelare. Nel caso del diritto d'autore applicato all'IA, questo significa chiedersi quali principi normativi – tutela della creatività umana, incentivo all'innovazione, accesso alla conoscenza, equa distribuzione dei benefici – debbano orientare le scelte del legislatore e in che modo tali principi possano essere bilanciati nel momento in cui entrano in conflitto.

Questa triplice dimensione – vale a dire: genealogica, analitico-concettuale e normativa – caratterizza l'approccio filosofico-giuridico e lo distingue, tanto, dalla mera analisi tecnico-giuridica, quanto, dalla speculazione filosofica astratta. Non si tratta né di applicare meccanicamente le norme vigenti ai nuovi fenomeni, né di elaborare teorie filosofiche senza preoccuparsi della loro traducibilità in soluzioni giuridiche concrete.

Si tratta, piuttosto, di muoversi costantemente tra teoria e prassi, tra astrazione concettuale e attenzione ai dettagli normativi, mantenendo sempre presente l'esigenza di individuare quali categorie giuridiche, fondate su quali presupposti filosofici, possono meglio rispondere alle sfide poste dall'IA in modo da realizzare i valori fondamentali che l'ordinamento è chiamato a tutelare.

In questa sede è quanto mai opportuno riflettere criticamente sui limiti e sulle potenzialità dell'analisi gius-filosofica applicata all'innovazione tecnologica. Ciò è necessario per evitare, sia, un'eccessiva fiducia nelle capaci-

<sup>11</sup> Per un inquadramento analitico dell'interpretazione e dell'argomentazione giuridica, cfr. G. TARELLO, *L'interpretazione della legge*, Milano, Giuffrè, 1980; R. GUASTINI, *Interpretare e argomentare*, Milano, Giuffrè, 2011.

tà della teoria di risolvere problemi pratici complessi, sia, uno scetticismo paralizzante che nega qualsiasi utilità alla riflessione filosofica in ambito giuridico. Si osservi, la consapevolezza dei limiti intrinseci dell'approccio filosofico-giuridico non ne sminuisce il valore, al contrario consente di utilizzarlo in modo ancor più efficace, integrandolo con altri strumenti di analisi e rafforzandolo quando appaia inadeguato o di per sé insufficiente.

Un primo limite evidente dell'analisi filosofica riguarda il suo carattere inevitabilmente astratto e generale. La filosofia lavora con concetti universali e cerca di individuare strutture invarianti al di là delle contingenze empiriche. Questa propensione all'astrazione, se da un lato permette di cogliere aspetti fondamentali che potrebbero sfuggire all'analisi empirica, dall'altro rischia di produrre teorie troppo evanescenti e distanti dalla realtà concreta per essere effettivamente utilizzate. Come ha osservato Niklas Luhmann nella sua critica alla filosofia del diritto tradizionale: "l'eccessiva distanza dal sistema giuridico operante condanna molte riflessioni filosofiche all'irrelevanza pratica"<sup>12</sup>. Nel caso dell'IA, ovviamente, questo rischio si fa particolarmente acuto: discutere in astratto se l'IA possa essere creativa può tradursi in una speculazione sterile se, parallelamente, non si considerano le specificità tecniche dei diversi sistemi, non si tiene conto delle modalità concrete del loro utilizzo e non si hanno presenti i contesti sociali in cui tali tecnologie operano.

Un secondo limite riguarda la dimensione temporale. L'innovazione tecnologica procede a una velocità che mal si concilia con i tempi lunghi della riflessione filosofica. Nel momento in cui l'analisi concettuale approfondita si esaurisce, la tecnologia sulla quale essa verteva potrebbe essere già stata sostituita da un ritrovato successivo, con caratteristiche completamente diverse. Questo sfasamento temporale suscita un problema metodologico e pratico molto serio, poiché potrebbe sembrare insensato elaborare categorie filosofiche pensate per rispondere a fenomeni *in re ipsa* transitori. E se è pur vero che certe questioni filosofiche mantengono la loro rilevanza al di là dei mutamenti contingenti, è altrettanto vero che la speculazione filosofica deve essere aggiornata rispetto agli sviluppi per non rischiare di indugiare e di discutere di problemi ormai superati.

Un terzo limite concerne la pluralità di prospettive filosofiche irriducibili. Come si vedrà nei paragrafi successivi, le diverse tradizioni filosofiche – aristotelica, cartesiana, empirista, fenomenologica, analitica – offrono definizioni e sistematizzazioni dell'intelligenza e della creatività non facilmente conciliabili. Questa pluralità può essere vista come una ricchezza

<sup>12</sup> N. LUHMANN, *Das Recht der Gesellschaft*, Frankfurt am Main, Suhrkamp, 1993.

(poiché permette di illuminare aspetti diversi della questione), ma può anche generare una paralisi decisionale: se non vi è accordo su che cosa sia l'intelligenza, come si può pretendere di elaborare soluzioni giuridiche univoche? La risposta a questo problema non può consistere nell'imporre una visione come l'unica corretta, ma, piuttosto, nel riconoscere che il diritto deve operare anche in condizioni di "dissenso ragionevole", elaborando soluzioni pragmatiche che non presuppongono la comune convergenza.

Un quarto limite riguarda il rapporto tra teoria e prassi. L'analisi filosofico-giuridica può individuare contraddizioni concettuali, evidenziare presupposti impliciti, proporre distinzioni analitiche, ma non può sostituirsi al lavoro concreto di elaborazione normativa, interpretazione giurisprudenziale e applicazione pratica. Come ha sottolineato Joseph Raz, "la filosofia del diritto non può dire ai giudici come decidere i casi concreti, ma può aiutarli a comprendere meglio la natura dell'attività interpretativa che stanno svolgendo"<sup>13</sup>. Nel contesto dell'IA e del diritto d'autore, questo significa che l'analisi filosofica può chiarire quali siano i valori in gioco e quali le opzioni siano concettualmente coerenti, ma non può determinare quale soluzione debba essere adottata nel caso specifico, poiché tale scelta dipende anche da considerazioni di opportunità politica, di fattibilità tecnica, di accettabilità sociale.

Nonostante questi limiti, l'analisi filosofico-giuridica possiede potenzialità significative che non soltanto giustificano ma impongono il suo impiego nello studio dell'innovazione tecnologica. Una prima potenzialità risiede nella capacità di anticipazione concettuale. Mentre l'analisi empirica deve attendere che i fenomeni si manifestino per studiarli e mentre il diritto positivo può reagire solo *ex post* ai problemi emersi, la riflessione filosofica può individuare in anticipo i nodi problematici che l'innovazione tecnologica solleverà. Interrogandosi sui presupposti delle categorie giuridiche esistenti, la filosofia può prevedere dove tali categorie entreranno in crisi e suggerire direzioni di riforma prima che la crisi diventi acuta. Nel caso dell'IA generativa, ad esempio, l'analisi filosofica aveva già evidenziato le tensioni tra concezione romantica dell'autore e possibilità di creazione automatizzata ben prima che i sistemi attuali raggiungessero livelli di sofisticazione tali da rendere il problema urgente.

Una seconda potenzialità consiste nella funzione critica. L'analisi filosofica permette di sottoporre a scrutinio critico soluzioni che potrebbero apparire ovvie o naturali, mostrandone i presupposti ideologici e le con-

<sup>13</sup> J. RAZ, *Between Authority and Interpretation: On the Theory of Law and Practical Reason*, Oxford, Oxford University Press, 2009.

sequenze problematiche. Quando, ad esempio, si propone di attribuire la titolarità delle opere generate da IA all'utente che ha fornito il *prompt*, l'analisi filosofica può interrogarsi su quale teoria dell'autorialità tale soluzione presuppone, se essa sia coerente con i principi generali del diritto d'autore e quali effetti essa produrrebbe sul lungo periodo. Questa funzione è essenziale per evitare soluzioni affrettate che, pur risolvendo problemi immediati, potrebbero creare contraddizioni sistemiche o violare principi fondamentali.

Una terza potenzialità risiede nella dimensione normativa. Come si è detto, l'analisi filosofica non si limita a descrivere o a criticare, ma propone criteri per valutare e orientare le scelte normative. Nel campo del diritto e della tecnologia, in cui gli interessi in gioco sono molteplici e spesso confliggenti, è essenziale disporre di una visione ontofenomenologica del diritto<sup>14</sup> e della giustizia che permetta di gerarchizzare i valori e di bilanciare le esigenze contrapposte. L'analisi filosofico-giuridica può fornire tale teoria, attingendo alla tradizione del pensiero politico e morale per individuare i principi che devono guidare la regolamentazione dell'innovazione: la dignità della persona, l'autonomia individuale, l'equa distribuzione dei benefici e degli oneri, la trasparenza e *accountability*, la tutela dei diritti fondamentali.

Una quarta potenzialità, infine, riguarda la funzione di integrazione concettuale. In un campo come quello dell'IA e del diritto, in cui convergono discipline eterogenee con linguaggi e metodi diversi, è necessario un lavoro di traduzione e armonizzazione che permetta il dialogo effettivo. La filosofia, per la sua vocazione alla generalità e alla chiarificazione concettuale, può svolgere questo ruolo di mediazione, individuando i punti di contatto tra prospettive diverse e costruendo ponti concettuali che facilitino la comunicazione interdisciplinare.

In conclusione, l'analisi filosofico-giuridica dell'innovazione tecnologica deve essere sì consapevole dei propri limiti – astrazione eccessiva, sfasamento temporale, pluralità di prospettive, distanza dalla prassi – ma anche conscia dei propri elementi di forza. La sua funzione non è quella di fornire soluzioni definitive ai problemi giuridici posti dall'IA, ma di chiarificare i termini del dibattito, di individuare le opzioni concettualmente coerenti, di evidenziare i valori in gioco e di proporre criteri normativi per orientare le scelte dei decisori. Si tratta della capacità di illuminare la complessità senza pretendere di eliminarla. Ed è proprio in questo aspetto che risiede il valore irrinunciabile dell'approccio filosofico-giuridico.

<sup>14</sup> Il rinvio, ovviamente, è a quella prospettiva che fu mirabilmente tracciata da S. COTTA (*Il diritto nell'esistenza. Linee di ontofenomenologia giuridica*, cit.).

### 3. *La creatività come concetto giuridicamente rilevante*

Se il concetto di intelligenza, come si è visto nel paragrafo precedente, rappresenta un problema ontologico fondamentale per l'ordinamento giuridico, la creatività costituisce il precipitato naturale di tale problema nell'ambito del diritto d'autore. Non è possibile, infatti, comprendere le sfide che l'IA generativa pone alla disciplina della proprietà intellettuale senza interrogarsi preliminarmente sulla natura della creatività e sul suo statuto epistemologico. Diventa, quindi, necessario chiarire anzitutto il significato stesso dell'affermazione in base alla quale un'azione è da ritenersi creativa. Nel dettaglio, occorre stabilire se la creatività sia una capacità esclusivamente umana o possa essere ascrivibile anche a sistemi artificiali. Risulta, inoltre, fondamentale individuare non solo criteri utili per l'attribuzione dei diritti ma anche per l'imputazione delle responsabilità. Tutte queste questioni attengono alla possibilità stessa di applicare le categorie tradizionali del diritto d'autore ai prodotti dell'IA. Come ha osservato Margaret Boden nella sua monumentale analisi della creatività computazionale "la creatività non è un dono misterioso conferito solo agli esseri umani, ma un insieme di capacità cognitive che possono essere analizzate, comprese e, in linea di principio, replicate artificialmente"<sup>15</sup>. Questa tesi, per quanto controversa, pone una sfida radicale all'impianto concettuale del diritto d'autore, che ha storicamente costruito la propria legittimazione sulla presunta unicità e irripetibilità dell'atto creativo umano.

L'indagine qui condotta si misura direttamente con tale sfida cercando di elaborare un quadro concettuale che permetta di valutare le opere generate dall'IA senza cadere né in un riduzionismo tecnologico che misconosca la specificità della creatività umana, né in un antropocentrismo acritico che rifiuti a priori di riconoscere valore alle produzioni artificiali. Per raggiungere tale obiettivo, è indispensabile ripercorrere la genealogia filosofica del concetto di creatività, comprendere come tale concetto si sia tradotto in categorie giuridiche attraverso la teoria romantica dell'autore e, infine, interrogarsi sulla possibilità di estendere o riformulare tali categorie alla luce dei fenomeni di creatività algoritmica. Solo attraverso questo percorso sarà possibile individuare i criteri normativi che devono ispirare la regolamentazione giuridica dell'IA generativa nell'ambito del diritto d'autore.

La posta in gioco – come è facile intuire – è particolarmente elevata. Il concetto di creatività non ha solo una dimensione descrittiva, ma svolge una

<sup>15</sup> M.A. BODEN, *The Creative Mind: Myths and Mechanisms* (2a ed.), London, Routledge, 2004.

funzione normativa centrale nel sistema del diritto d'autore. È in ragione della creatività che si giustifica l'attribuzione all'autore di un diritto esclusivo di sfruttamento economico dell'opera ed è sempre in ragione della creatività che si distinguono le opere proteggibili dalle mere produzioni meccaniche o triviali.

Nel momento in cui l'IA è in grado di produrre artefatti creativi in senso proprio, allora il sistema giuridico deve decidere se estendere la protezione autorale a tali artefatti e, in caso affermativo, a chi attribuire la titolarità dei relativi diritti. Se, al contrario, si ritiene che la creatività presupponga caratteristiche specificamente umane che l'IA non possiede, allora occorre individuare con precisione quali siano tali caratteristiche e perché esse risultino giuridicamente decisive. In entrambi i casi, è necessaria una riflessione filosofica profonda, che vada oltre le approssimazioni intuitive e costruisca un apparato concettuale rigoroso.

La nozione di creatività è il prodotto di una lunga evoluzione storica e filosofica che ha attraversato profonde trasformazioni semantiche e assiologiche. Nella Grecia antica, la creazione artistica non era concepita come l'esternazione e il frutto di un potere autonomo dell'individuo, ma come risultato di un'ispirazione (*enthousiasmós*) di origine divina. Nello *Ione*, Platone descrive il poeta come posseduto dalle Muse, che lo rendono strumento passivo della rivelazione del bello e del vero: "il poeta è cosa leggera, alata e sacra, e non è in grado di comporre prima di essere ispirato dal dio e di uscire fuori di sé e di non avere più intelletto"<sup>16</sup>. Se così, è evidente che la creatività non appartiene all'essere umano, ma al divino, che opera attraverso di lui. L'artista non crea in senso proprio, bensì riceve e trasmette.

Tale concezione teurgica della creatività ha dominato per secoli la cultura occidentale, trovando nuova legittimazione nella teologia cristiana, in cui la *creatio ex nihilo* era un attributo esclusivo di Dio. San Tommaso d'Aquino – riprendendo e sviluppando il pensiero aristotelico – distingue nettamente la creazione divina (in grado di produrre l'essere dal nulla), dalla produzione artistica umana (che si limita a dare forma a una materia preesistente)<sup>17</sup>. L'artista, in questa prospettiva scolastica, non è creatore in senso proprio, ma artefice che imita le forme naturali impresse da Dio nella creazione. La sua attività, per quanto meritoria, rimane subordinata e derivata rispetto all'unico atto creativo autentico, quello divino, poiché per Tommaso l'uomo opera per *modum artis* (secondo l'arte), mentre Dio per *modum naturae* (creando la natura stessa).

<sup>16</sup> PLATONE, *Ione*, in ID., *Tutti gli scritti*, Milano, Bompiani, 2000.

<sup>17</sup> San T. D'AQUINO, *La Somma Teologica*, Vol. I, Bologna, Edizioni Studio Domenicano, 1988.

La svolta decisiva nel segno di una concezione antropocentrica della creatività si registra nel Rinascimento. È Leon Battista Alberti, nel suo trattato *De pictura*<sup>18</sup>, a porre le basi per l'idea dell'artista come *alter deus*, capace – non soltanto di imitare passivamente la natura – ma anche di emularla e, persino, di superarla attraverso l'opera di ingegno. Questa rivendicazione della dignità creatrice dell'essere umano si accompagna a una progressiva secolarizzazione del concetto di genio, che – da prerogativa e da dono divino – diventa facoltà naturale dell'individuo eccezionale<sup>19</sup>.

Questo processo troverà poi piena espressione nel neoplatonismo di Ficino, che propone, infatti, una mediazione tra l'antica concezione teurgica e la nuova visione antropocentrica: il genio è una forza divina che opera nell'anima umana, ma lo fa stimolando le potenzialità creative già insite nell'individuo. È, tuttavia, con l'Illuminismo e il Romanticismo che la creatività assume il significato di capacità propriamente umana di produrre opere originali attraverso l'immaginazione e il talento individuale. Kant, nella sua *Critica del Giudizio*, elabora una teoria sistematica del genio come “talento (dono naturale) che dà la regola all'arte”<sup>20</sup>. Il genio – a detta del filosofo di Königsberg – non applica regole pre-esistenti, ma crea regole nuove attraverso le quali la natura stessa si esprime nell'arte. Questa concezione segna una rottura definitiva rispetto alla precedente visione sostanzialmente (e meramente) imitativa dell'arte: il creatore, d'ora in poi, non riproduce più ciò che già esiste, ma porta all'esistenza qualcosa di radicalmente nuovo che non può essere dedotto da principi antecedenti. La creatività diventa, così, la capacità di generare *ex novo*, prospettando un qualcosa che non è riducibile alla mera combinazione meccanica di elementi già dati. Il Romanticismo porta alle estreme conseguenze questa valorizzazione della creatività come attributo fondamentale dell'umano. Per Friedrich Schelling, l'attività artistica costituisce la forma più alta di conoscenza, superiore sia alla scienza che alla filosofia, perché in essa l'infinito si manifesta nel finito e l'assoluto diventa intuibile sensibilmente<sup>21</sup>. L'artista

<sup>18</sup> L.B. ALBERTI, *De pictura* (1435), in ID., *Opere volgari*, Bari, Laterza, 1973. La formula dell'*alter deus* va intesa qui come lettura interpretativa del ruolo dell'artista nella riflessione umanistica, più che come espressione testuale letterale di Alberti.

<sup>19</sup> Questo processo troverà poi piena espressione nel neoplatonismo di Ficino e nella trattatistica cinquecentesca. Marsilio Ficino, nella sua filosofia neoplatonica, propone infatti una mediazione tra l'antica concezione teurgica e la nuova visione antropocentrica: il genio è una forza divina che opera nell'anima umana, ma lo fa stimolando le potenzialità creative già insite nell'individuo. Cfr. M. FICINO, *De amore* (1469), in ID., *El libro dell'Amore*, Firenze, Olschki, 1987.

<sup>20</sup> I. KANT, *Critica della capacità di giudizio*, trad. it., Milano, BUR, 2004.

<sup>21</sup> F.W.J. SCHELLING, *Sistema dell'idealismo trascendentale*, trad. it., Milano, Bompiani, 1989.

non è più né artigiano né imitatore, ma creatore originario che rivela dimensioni altrimenti inaccessibili della realtà. Ancora, Samuel Taylor Coleridge distingue tra fantasia (*fancy*) che si limita a combinare elementi dati dalla memoria e immaginazione (*imagination*) che è “il potere vivente e primo di ogni percezione umana” e “ripetizione nell’anima finita dell’eterno atto della creazione nell’Io infinito”<sup>22</sup>. In questa prospettiva, la creatività autentica partecipa della stessa potenza creatrice divina, pur esercitandosi nell’ambito finito dell’esperienza umana.

Nel corso del Novecento, le scienze cognitive e la psicologia sperimentale hanno tentato di demistificare il concetto di creatività, analizzandola come insieme di processi cognitivi studiabili empiricamente. Joy Paul Guilford – pioniere della ricerca psicologica sulla creatività – ha proposto una distinzione tra pensiero convergente che ricerca la soluzione unica e corretta a un problema ben definito e pensiero divergente che genera molteplici soluzioni originali ed eterodosse<sup>23</sup>. Questa concettualizzazione ha avuto il merito di elevare la creatività alla stregua di un oggetto di indagine scientifica, ma ha anche suscitato aspre critiche per via del suo carattere riduttivo: limitare la creatività all’elemento della fluidità, della flessibilità e dell’originalità del pensiero rischia di far perdere di vista la dimensione valoriale e culturale che, storicamente, ha connotato tale concetto.

Più recentemente, la filosofia della creatività ha cercato di elaborare modelli che integrino la dimensione cognitiva con quella sociale e culturale. In modo particolare, Margaret Boden ha proposto una distinzione tra creatività psicologica (*P-creativity*), che produce idee nuove per il soggetto che le concepisce, e creatività storica (*H-creativity*), che genera idee nuove per l’intera umanità o per un campo culturale specifico<sup>24</sup>. Questa distinzione permette di riconoscere che la creatività non è semplicemente una proprietà intrinseca di un’opera o di un processo mentale, ma dipende dal contesto storico e dalla tradizione entro cui si colloca. Un’idea può essere *P-creativa* per chi la concepisce anche se è già stata formulata da altri, ma solo se è genuinamente *H-creativa* può avere rilevanza culturale e, di riflesso, giuridica.

Sul piano del diritto d’autore, questa genealogia filosofica ha una serie di conseguenze dirette. La concezione moderna della creatività – come capacità umana autonoma di generare novità – ha fornito il fondamento teorico per l’attribuzione all’autore di diritti esclusivi sull’opera. Se l’autore è creatore nel senso “forte” del termine, se porta a esistenza qualcosa

<sup>22</sup> S.T. COLERIDGE, *Biographia Literaria*, Princeton, Princeton University Press, 1983.

<sup>23</sup> J.P. GUILFORD, “Creativity”, in *American Psychologist*, 5 (9), 1950, pp. 444-454.

<sup>24</sup> M.A. BODEN, *The Creative Mind: Myths and Mechanisms*, London, Routledge, 2004.

che prima non c'era attraverso l'esercizio delle proprie facoltà individuali, allora appare legittimo riconoscergli un diritto di proprietà sulla sua stessa creazione. Tuttavia, come si vedrà nel prosieguo, questa legittimazione filosofica della proprietà intellettuale presuppone una concezione della creatività che l'IA mette radicalmente in discussione. Se un algoritmo può generare opere originali senza possedere le caratteristiche che la tradizione filosofica ha sempre associato alla creatività umana – vale a dire, coscienza, intenzionalità, esperienza vissuta – allora è evidente che occorre ripensare i fondamenti stessi del diritto d'autore o elaborare categorie concettuali nuove capaci di dar conto di questa inedita situazione.

La concezione romantica della creatività, analizzata nel paragrafo precedente, ha trovato una traduzione giuridica ben precisa nella teoria dell'autore, che ha segnato l'intero sistema moderno della tutela diritto d'autore. Questa teoria, sviluppatasi principalmente nel corso del XIX secolo, identifica l'autore come soggetto dotato di personalità creativa, capace di imprimere nell'opera la propria individualità e meritevole, per tale ragione, di una tutela giuridica speciale. Come ha osservato Peter Jaszi, nel suo celebre saggio sulla costruzione dell'autorialità, “la figura romantica dell'autore solitario ispirato dal genio è diventata non solo un mito culturale dominante, ma anche il fondamento delle dottrine giuridiche del *copyright*”<sup>25</sup>.

Ovviamente, comprendere i presupposti filosofici di tale costruzione è essenziale per valutare se – e se sì in quale misura – essa possa essere estesa alle creazioni generate dall'IA. Il nucleo teorico di tale concezione risiede nell'idea che l'opera sia emanazione della personalità del creatore, alla stregua di un'espressione “esternalizzata” e “concretizzata” della sua interiorità spirituale. Questa tesi trova una delle sue formulazioni più compiute nella filosofia del diritto di Hegel, il quale – nella sezione dei *Lineamenti di filosofia del diritto* dedicata alla proprietà intellettuale – afferma che l'autore ha un diritto di proprietà sull'opera, in quanto essa costituisce un'estensione della sua personalità nel mondo esterno. L'opera letteraria o artistica non è un semplice oggetto materiale, ma “una cosa che, in quanto esiste esternamente, è in modo del tutto proprio mia cosa”<sup>26</sup>. Altrimenti detto ed in breve, la protezione giuridica dell'opera, in questa prospettiva, trova fondamento e giustificazione nella necessità di tutelare l'autonomia e la dignità della persona che in questa creazione in un certo senso si “oggettiva”.

<sup>25</sup> P. JASZI, “Toward a Theory of Copyright: The Metamorphoses of ‘Authorship’”, in *Duke Law Journal*, 1991(2), 1991, pp. 455-502; M. ROSE, *Authors and Owners: The Invention of Copyright*, Cambridge (MA), Harvard University Press, 1993.

<sup>26</sup> G.W.F. HEGEL, *Lineamenti di filosofia del diritto*, trad. it., Roma-Bari, Laterza, 1996.

Tale impostazione ha esercitato un'influenza profonda sulla tradizione giuridica continentale, in particolare sul sistema francese del *droit d'auteur* e tedesco dell'*Urheberrecht*, che riconoscono all'autore, non solo diritti patrimoniali, ma anche diritti morali inalienabili sull'opera. Nel dettaglio, il diritto morale d'autore – che comprende il diritto di paternità, il diritto all'integrità dell'opera, il diritto di divulgazione e il diritto di pentimento – presuppone che esista un legame personale e indissolubile tra l'autore e la sua creazione, un legame che la mera cessione dei diritti economici non può comunque sia recidere. Come è stato affermato dalla giurisprudenza francese in numerose occasioni, "l'opera è il prolungamento della personalità del suo autore, ed essa deve quindi godere di una protezione che trascende i meri interessi economici"<sup>27</sup>.

Anche nel sistema anglosassone del *copyright* – tradizionalmente più orientato agli aspetti economici che ai diritti morali – la teoria romantica dell'autore ha lasciato alcune tracce significative. Sebbene la giustificazione utilitaristica della proprietà intellettuale – in base alla quale il *copyright* servirebbe a incentivare la produzione di opere socialmente utili – sia stata prevalente negli Stati Uniti e nel Regno Unito, la dottrina dell'originalità (come manifestazione della personalità dell'autore) ha comunque trovato spazio. Significativa, in tal senso, la sentenza *Feist Publications v. Rural Telephone Service Co.* – pronunciata dalla Corte Suprema degli Stati Uniti – nella quale è stato stabilito che "l'originalità richiesta per la protezione del *copyright* non è semplicemente novità, ma presuppone un minimo di creatività", intendendo con ciò una precisa e inedita scelta personale che rifletta la visione individuale dell'autore<sup>28</sup>.

La teoria romantica dell'autore ha però anche attirato diverse critiche sostanziali e ciò, sia sul piano filosofico sia su quello sociologico. Ad esempio, Michel Foucault – nel celebre saggio *Qu'est-ce qu'un auteur?* – ha decostruito la figura dell'autore mostrandone la natura storicamente contingente e funzionale<sup>29</sup>. In modo particolare, l'autore, secondo Foucault, non è un'entità naturale pre-esistente rispetto all'opera, ma incarna più che altro una sorta di funzione-autore che è stata costruita *ad hoc* dal discorso giuridico e dall'elaborazione critica per conferire unità e coerenza a un insieme di testi. Sempre in chiave critica, anche Roland Bar-

<sup>27</sup> *Cour de Cassation* (Francia), sentenza del 28 maggio 1991, *Huston c. Société Turner Entertainment*, in *Revue internationale du droit d'auteur*, n. 149, p. 197.

<sup>28</sup> *Feist Publications, Inc. v. Rural Telephone Service Co.*, United States Supreme Court, 499 U.S. 340 (1991).

<sup>29</sup> M. FOUCAULT, "Qu'est-ce qu'un auteur?", in *Bulletin de la Société Française de Philosophie*, 63(3), 1969, pp. 73-104.

thes, che – con una visione è ancor più radicale – ha proclamato la morte dell'autore, sostenendo che il significato di un'opera non risiede nell'intenzione originaria del creatore, bensì nelle molteplici interpretazioni del lettore<sup>30</sup>. Ed è interessante notare che, per Barthes, il testo può essere visto come una specie di “tessuto di citazioni”. Una lettura che, oggi, può essere richiamata in via analogica anche con riguardo al funzionamento dei *large language models*, nella misura in cui essi si fondano sulla rielaborazione statistica di vasti corpora testuali di addestramento. Vien da sé che queste critiche post-strutturaliste – pur non negando l'utilità pratica della categoria giuridica di autore – ne hanno però evidenziato il carattere ideologico e la funzione di “occultamento” delle pratiche collaborative e intertestuali che caratterizzano la produzione culturale.

Ora, dal punto di vista del diritto d'autore contemporaneo, la persistenza dell'influsso della teoria romantica dell'autore si manifesta in almeno tre ambiti fondamentali: 1) nel “requisito dell'originalità” in ragione della quale un'opera sarebbe “proteggibile” se dotata di carattere creativo, ovvero sia di un'impronta della personalità dell'autore (questa è la definizione fornita dalla Corte di Giustizia dell'Unione Europea (sentenze Infopaq<sup>31</sup> e Painer<sup>32</sup>); 2) nella “presunzione di titolarità”, secondo cui l'autore è colui che ha creato l'opera e, salvo disposizioni specifiche (come nel caso delle opere collettive o delle opere realizzate su commissione), è il titolare originario dei diritti; 3) nei “diritti morali”, che si declinano nella possibilità di rivendicare la paternità dell'opera e di opporsi a modifiche che ledano la reputazione dell'autore in virtù di quel peculiare legame personale che soltanto la persona può intrattenere con le proprie creazioni.

L'avvento dell'IA generativa, come è intuitivo, mette in crisi ciascuno di questi tre pilastri. Infatti, nel momento in cui un algoritmo può generare testi, immagini o musiche che manifestano caratteristiche formalmente indistinguibili da quelle delle opere umane, affiora la necessità (e la difficoltà) di comprendere in che senso si possa ancora parlare di “impronta della personalità”. Emerge anche un ulteriore aspetto di complessità circa la necessità di identificare l'autore di un'opera generata dall'IA, in ragione delle diverse figure coinvolte: i) il programmatore del sistema; ii) l'utilizzatore che ha fornito il *prompt*; iii) l'azienda proprietaria dell'algoritmo; iv) nessuno di essi. Diventa, infine, necessario stabilire se abbia

<sup>30</sup> R. BARTHES, “La mort de l'auteur”, in *Manteia*, 5, 1967, pp. 12-17.

<sup>31</sup> CGUE, sentenza del 16 luglio 2009, causa C-5/08, *Infopaq International A/S v. Danske Dagblades Forening*, EU:C:2009:465.

<sup>32</sup> CGUE, sentenza del 1 dicembre 2011, causa C-145/10, *Eva-Maria Painer v. Standard VerlagsGmbH and Others*, EU:C:2011:798.

ancora senso parlare di diritti morali in relazione a opere che non sono frutto dell'ingegno di un creatore umano identificabile.

In questa sede si intende affrontare questa sfida, cercando di individuare quali elementi della teoria romantica dell'autore possano essere preservati in un contesto di creatività ibrida uomo-macchina e quali aspetti debbano invece essere rivisti o sostituiti. L'obiettivo non è quello di difendere a ogni costo una concezione ormai superata, ma nemmeno quello di abbandonare precipitosamente categorie che hanno dimostrato la propria utilità nel corso dei secoli e che – se risemantizzate, adeguate e aggiornate – potrebbero rivelarsi tuttora utili. In modo particolare, si tratta, per un verso, di comprendere quali sono le funzioni svolte dalla figura dell'autore all'interno del sistema del diritto d'autore e, per un altro verso, di verificare se tali funzioni possano essere assolte in modi diversi nel momento in cui l'autore umano cede il posto (in tutto o in parte) a creatori macchinici e sistemi artificiali. Solo in questo modo sarà possibile elaborare soluzioni normative formalmente accettabili e sostanzialmente auspicabili, poiché, al contempo, eticamente fondate e tecnologicamente consapevoli.

Il dibattito circa la possibilità (o meno) di riconoscere all'IA l'attributo della creatività (in senso proprio) individua una *vexata quaestio* e rappresenta uno degli snodi teorici più ostici nel dibattito contemporaneo. Vien da sé che se si ammette che l'IA possa essere creativa, allora diventa assai difficile negare *a priori* che le sue produzioni meritino una qualche forma di tutela; se, al contrario, si sostiene che la creatività presuppone caratteristiche specificamente e squisitamente umane, allora è evidente che le opere generate dall'IA dovrebbero essere escluse *in re ipsa* dalla protezione del diritto d'autore.

Una prima interessante posizione, è stata avanzata da Margaret Boden e da David Cope ed afferma che la creatività algoritmica rappresenta sostanzialmente una sorta di *continuum* rispetto alla creatività umana. Più nel dettaglio, a detta di Boden e Cope, la creatività non dovrebbe essere ritenuta alla stregua di una capacità misteriosa e/o trascendente, in quanto essa è più che altro il risultato di un insieme di processi cognitivi – di combinazione, di esplorazione e di trasformazione – che possono anche essere replicati computazionalmente. Da par suo, Boden prospetta e distingue tre fondamentali forme di creatività: 1) quella combinatoria; 2) quella esplorativa; 3) quella trasformazionale<sup>33</sup>. La creatività combinatoria genera idee nuove associando elementi già esistenti in modi inediti;

<sup>33</sup> M.A. BODEN, *op. cit.*

la creatività esplorativa esplora sistematicamente uno spazio concettuale predefinito alla ricerca di soluzioni originali; infine, la creatività trasformativa modifica le regole stesse dello spazio concettuale, aprendo possibilità prima impensabili. Tutte e tre queste forme di creatività – secondo Boden – possono essere implementate anche algoritmicamente, come è dimostrato dai sistemi di IA che compongono musica “nello stile di Bach”, scrivono poesie “con il linguaggio di D’Annunzio” o generano immagini artistiche che “sembrano un prodotto del genio di Van Gogh”.

A questo proposito non si può non ricordare che David Cope – compositore e informatico – ha sviluppato il programma EMI (*Experiments in Musical Intelligence*), concepito per analizzare lo stile di diversi compositori classici e generare nuove composizioni che, in alcuni esperimenti e test comparativi, sono state giudicate da parte del pubblico o di ascoltatori non specialisti difficilmente distinguibili da quelle originali, alimentando così un intenso dibattito circa la possibilità di ritenerle creative. Cope sostiene che “*se accettiamo che la creatività consiste nel generare strutture nuove rispettando certi vincoli stilistici, allora non vi è ragione per negare che EMI sia creativo*”<sup>34</sup>. Qui è utile precisare che Cope definisce la creatività in termini di *recombinantics* (ricombinazione), ovvero attraverso un termine tecnico che usa spesso per spiegare come EMI riutilizzi segmenti di opere esistenti. Questa argomentazione, però, presuppone una definizione funzionale e comportamentale della creatività, in base alla quale è creativo ciò che produce risultati originali secondo criteri condivisi, indipendentemente dal processo sottostante.

Una posizione diametralmente opposta è stata sostenuta da Hubert Dreyfus e da John Searle, che, di fatto, negano che l’IA possa essere ritenuta genuinamente creativa, in quanto è priva di quella comprensione intenzionale e di quell’esperienza vissuta che costituiscono il presupposto distintivo ed imprescindibile della creatività umana. Nel dettaglio, secondo Searle – che prospetta il celebre esperimento mentale della *Stanza cinese* – un sistema che manipola simboli secondo regole sintattiche ma pur senza comprenderne il significato semantico non può essere considerato intelligente in senso proprio<sup>35</sup>. Declinato in tema di creatività, questo stesso argomento suggerirebbe che un algoritmo che genera opere artistiche combinando *pattern* statistici estratti da un *corpus* di addestramento

<sup>34</sup> D. COPE, *Virtual Music: Computer Synthesis of Musical Style*, Cambridge (MA), MIT Press, 2001.

<sup>35</sup> J.R. SEARLE, “Minds, Brains, and Programs”, in *Behavioral and Brain Sciences*, 3(3), 1980, pp. 417-424.

“non sta creando” nel senso autentico del termine, ma sta semplicemente eseguendo operazioni meccaniche che producono *output* formalmente simili a opere creative senza però possederne la dimensione intenzionale.

Una terza posizione – che può essere definita intermedia o pluralista – sostiene che la creatività algoritmica non sia né riducibile alla creatività umana né “totalmente altra” rispetto ad essa, ma integri una forma emergente di creatività che richiede categorie concettuali nuove. Di questo avviso è, ad esempio, Aaron Sloman. Filosofo dell’intelligenza artificiale, Sloman ha proposto di dismettere la domanda binaria: *l’IA può essere creativa?* Egli auspica, infatti, un’analisi più articolata dei diversi tipi di creatività e dei diversi livelli di comprensione che possono caratterizzarli<sup>36</sup>. Si badi, però, Sloman non parla di intelligenza al singolare, quanto piuttosto di una architettura della mente (modello *H-CogAff*) e tale assunto comporterebbe il fatto che l’IA possa possedere una “comprensione funzionale” (cioè la capacità di manipolare relazioni tra concetti) anche se priva di “comprensione fenomenica”, vale a dire della dimensione soggettiva qualitativa dell’esperienza – i cosiddetti *qualia* – che caratterizza il sentire consapevole. Elemento, che giustificherebbe una tutela giuridica basata sulla funzione dell’opera. Se così, alcune forme di creatività computazionale potrebbero essere considerate “autentiche” anche se non replicano esattamente i processi della creatività umana, purché soddisfino dei criteri appropriati di novità, di valore e di autonomia.

Si osservi, la ripresa – e se vogliamo anche l’integrazione – delle intuizioni di Sloman nell’ambito del quadro teorico qui proposto non si ritraduce nell’adesione a un realismo ingenuo che considera la creatività alla stregua di una proprietà fisica della macchina. Il punto è un altro, in quanto l’analisi dei diversi livelli di “architettura mentale” proposta da Sloman si rivela propedeutica a dar conto della complessità del sistema. E, chiaramente, una volta che tale complessità è stata tecnicamente accertata, può innestarsi la – diversa ma necessaria – prospettiva di Searle utile a operare il salto normativo: il diritto, invero, non “scopre” una creatività pre-esistente nell’algoritmo, ma decide di “attribuire” convenzionalmente lo *status* di creatività a determinati *output* di alto livello, trasformando un “fatto bruto” (quale è l’elaborazione dei dati) in un fatto sociale (ovverosia l’opera protetta) per finalità di utilità collettiva.

<sup>36</sup> A. SLOMAN, “Some Requirements for Human-Like Robots: Why the Recent Over-Emphasis on Embodiment Has Held Up Progress”, in B. SENDHOFF, E. KÖRNER, O. SPORNS, H. RITTER, K. DOYA (eds.), *Creating Brain-Like Intelligence*, Berlin, Springer, 2009, pp. 248-277.

Come è naturale, questa pluralità di posizioni teoriche si riflette in altrettanti approcci normativi differenti. Del resto: se si adotta la tesi della continuità, diventa plausibile estendere la protezione del diritto d'autore alle opere generate dall'IA, eventualmente attribuendo la titolarità del prodotto all'utilizzatore o al programmatore del sistema; se, invece, si adotta la tesi della discontinuità radicale, allora le opere generate dall'IA dovrebbero ricadere nel dominio pubblico, essendo prive di quel requisito di creatività personale che ne giustifica la particolare protezione; infine, se si adotta una posizione c.d. intermedia, potrebbe essere opportuno elaborare forme di tutela *ad hoc* che riconoscano valore alle produzioni dell'IA senza, però, equipararle in tutto e per tutto a quelle umane.

La via preferibile è quella di un approccio graduale e funzionale. In pratica, anziché decidere in astratto se l'IA possa o meno essere creativa, ci si muove verso una valutazione del singolo caso (operando una sorta di controllo creativo). L'idea è quella di considerare di volta in volta – e *secundum quid* – quale sia il grado di autonomia e di originalità mostrato da un sistema specifico, quale ruolo abbia giocato l'intervento umano nel processo generativo e quali interessi siano meritevoli di tutela.

Intanto, una precisazione tutt'altro che peregrina: un'opera generata interamente da un sistema autonomo suscita una serie di questioni molto diverse da un'opera prodotta con l'assistenza dell'IA ma pur sempre sotto la guida creativa dell'agente umano. E ancora: un algoritmo addestrato su un *corpus* di opere protette solleva problemi nettamente differenti rispetto a un sistema che genera contenuti *ex novo* senza utilizzare materiale coperto da *copyright*. L'approccio che si è scelto di adottare mira ad identificare il *creative nexus* (il nesso creativo) tra l'*input* umano (*prompting*, selezione, *editing*) e l'*output* prospettato dalla macchina. Infatti, nel momento in cui l'uomo mantiene il controllo sulle scelte espressive, la teoria romantica dell'autore può essere in qualche misura “salvata” tramite una reinterpretazione del ruolo della persona come “regista dell'algoritmo”.

Il punto è che questa prospettiva pragmatico-contestuale non risolve in radice le questioni filosofiche sottostanti, nonostante ciò, comunque, permette di affrontare i problemi giuridici concreti senza rimanere “bloccati” in questioni e in dispute teoriche pressoché irrisolvibili. Essa presuppone, però, una chiarificazione preliminare in ordine a quali siano i criteri normativi che devono guidare l'attribuzione di tutela. Laddove si aderisca a quella lettura secondo la quale il diritto d'autore è propedeutico a incentivare la produzione culturale, la questione diventa se la protezione delle opere generate dall'IA favorisca oppure ostacoli tale obiettivo. Nel caso in cui, invece, si interpreti il diritto d'autore come una forma

di tutela orientata ad assicurare l'autonomia creativa della persona, allora occorre chiedersi se – e se sì in quale misura – le opere generate dall'IA esprimano l'autonomia di qualcuno. Da ultimo, se si intravede nel diritto d'autore uno strumento di bilanciamento degli interessi dei creatori con quelli del pubblico, allora bisogna valutare quale regime normativo realizzi il migliore equilibrio.

Come si dirà nel prosieguo della trattazione, nessuna di queste domande tollera risposte semplici o definitive. Tuttavia, l'analisi filosofica della creatività algoritmica fornisce quegli strumenti concettuali necessari per affrontare tutti i quesiti e le varie istanze in modo rigoroso e consapevole. La sfida principale è costituita dalla necessità di riuscire a trasformare queste riflessioni teoriche in orientamenti normativi capaci di guidare i legislatori, i giudici e gli operatori del diritto nell'approccio all'IA generativa. Invero, solo grazie a un dialogo costante tra la filosofia – e in modo particolare la filosofia del diritto – e la dimensione pratico-normativa sarà possibile elaborare soluzioni che siano, al contempo, eticamente fondate, tecnologicamente informate e – soprattutto – giuridicamente auspicabili (perché improntate alla tutela della persona e dei suoi diritti fondamentali), oltre che concretamente e quotidianamente praticabili<sup>37</sup>.

#### 4. *Il problema ontologico dell'intelligenza non umana nell'ordinamento giuridico*

L'avvento e la rapidissima diffusione dei sistemi di IA generativa suscita un'istanza di natura ontologica, relativa alla possibilità di concepire l'intelligenza come una facoltà riconoscibile ad un'entità impersonale e alle conseguenze che tale eventuale riconoscimento può comportare in chiave giuridica. Come è immediatamente intuitivo, si tratta di una questione che investe i presupposti stessi sui quali si è edificato l'ordinamento giuridico moderno, che orbita attorno alla centralità del soggetto razionale (di matrice kantiana). La crisi che l'IA produce nel diritto d'autore, da questo punto di vista, costituisce un caso paradigmatico che può aprire a un ventaglio di riflessioni ulteriori e più ampie (che interessano categorie giuridiche classiche storicamente determinate), che nascono dalla intrinseca attitudine, che i fenomeni tecnologici hanno, di mettere

<sup>37</sup> In tal senso, impossibile non ricordare quanto – seppur in altro contesto – osservavano S. COTTA (*Il diritto nell'esistenza*, cit.; ID., *Diritto, persona, mondo umano*, cit.) e F. D'AGOSTINO (*Filosofia del diritto*, Torino, Giappichelli, 2006) a proposito dello scopo del diritto.

in discussione i fondamenti antropocentrici di alcuni istituti giuridici e di talune procedure.

In modo particolare, l'obiettivo è quello di armonizzare la tutela della proprietà intellettuale mediante un utilizzo responsabile dei sistemi di IA. Tale proposito, tuttavia, non può prescindere da una preliminare chiarificazione concettuale circa il significato della nozione di intelligenza nel momento in cui tale termine venga associato a un artefatto computazionale. In questa prospettiva, la filosofia della mente offre degli spunti decisivi per superare quell'annoso confinamento biologico della nozione di cognizione. Hilary Putnam, con il suo esternalismo semantico, ha avuto il merito di scardinare l'idea che i significati fossero fenomeni puramente interni ("*meanings just ain't in the head*"), sostenendo che il contenuto dei nostri stati mentali è parzialmente costituito dal mondo esterno<sup>38</sup>. Nonostante ciò, sono stati Andy Clark e David Chalmers a proporre le osservazioni più radicali e incisive per l'approccio allo studio dell'architettura della mente. Nel dettaglio, per mezzo della c.d. teoria della mente estesa, Clark e Chalmers hanno dimostrato che i processi cognitivi non sono vincolati e non si arrestano ai confini del cranio o della pelle, ma possono legittimamente incorporare supporti tecnologici esterni – dal semplice taccuino ai più articolati e complessi algoritmi di IA – a condizione che tali artefatti svolgano un ruolo funzionale integrato nel sistema<sup>39</sup>.

Nelle more del contesto giuridico, una siffatta revisione concettuale si rivela ancor più pregnante e multifaccettata: invero, se la mente è un sistema distribuito, accoppiato (e accoppiabile) con la tecnica, l'attribuzione di originalità e di autorialità non può più restare ancorata a una concezione strettamente psico-fisica e puramente interna della creatività umana, ma deve necessariamente aprirsi a una valutazione relazionale dell'*agency*.

Vien da sé – e non potrebbe essere altrimenti – che la questione si declini con accenti e con una urgenza del tutto particolare nel momento in cui si guarda al diritto d'autore: disciplina che più di ogni altra – come s'è detto – ha costruito il proprio impianto concettuale sull'idea di un soggetto creatore dotato di capacità intellettive umane.

Si osservi: la stessa originalità dell'opera – quale criterio cardine della protezione autorale – è stata definita dalla Corte di Giustizia dell'Unione Europea come espressione di una creazione intellettuale propria dell'autore, richiedendo che quest'ultimo abbia espresso scelte libere e creative tali

<sup>38</sup> H. PUTNAM, "Psychological Predicates", in W.H. CAPITAN, D.D. MERRIL (eds.), *Art, Mind, and Religion*, Pittsburgh, University of Pittsburgh Press, 1967, pp. 37-48.

<sup>39</sup> A. CLARK, D.J. CHALMERS, "The Extended Mind", in *Analysis*, 58(1), 1998, pp. 7-19.

da imprimere all'opera il proprio "tocco personale" e univoco (sentenza *Infopaq International* e sentenza *Painer*). Tale *standard* giurisprudenziale – ormai consolidato quale criterio generale di originalità nel diritto dell'Unione europea – presuppone manifestamente la presenza di un autore umano capace di autodeterminazione creativa. Il punto è che nel momento in cui l'IA produce artefatti che mostrano caratteristiche formalmente indistinguibili da quelle rintracciabili all'interno delle opere umane, il sistema giuridico si trova a misurarsi con un dilemma: o nega che tali prodotti possano integrare delle opere nel senso giuridicamente rilevante, preservando così la coerenza del paradigma antropocentrico ma escludendo dalla protezione una crescente fetta di creazioni; oppure si ammette che l'opera d'ingegno possa anche prescindere dall'autore umano, aprendo però il varco, in questo modo, a tutte le conseguenze che possono derivare sul piano della titolarità dei diritti e della responsabilità.

Vi è, poi, un ulteriore profilo di criticità – che affiora soprattutto agli occhi del filosofo del diritto – poiché non si tratta soltanto di decidere se l'IA sia o non sia intelligenza in senso proprio<sup>40</sup>, al contrario, si tratta di interrogarsi su quale modello di intelligenza risulti più adeguato a dar conto, insieme, delle capacità effettive dei sistemi computazionali e delle esigenze di tutela che l'ordinamento giuridico è chiamato a garantire. Da questo punto di vista, la questione ontologica dell'intelligenza non umana non può essere risolta né mediante un semplice rifiuto nominalistico – che neghi rilevanza giuridica ai sistemi artificiali in quanto privi di interiorità cosciente –, né mediante un'estensione acritica delle categorie tradizionalmente riservate al soggetto umano. L'esigenza dell'ordinamento non è quella di stabilire se la macchina sia "persona" o "mente" nello stesso senso in cui lo è l'essere umano, ma di comprendere se e in quale misura la sua attività debba essere giuridicamente qualificata in funzione degli effetti che produce, dei rapporti che struttura e degli interessi che coinvolge. In tale prospettiva, il problema ontologico non precede il diritto come un dato neutro da registrare, ma si intreccia con esso, poiché è il diritto stesso a dover selezionare, entro la pluralità dei modelli teorici disponibili, quelli che consentono di costruire categorie normative coerenti con la dignità della persona, con la responsabilità degli attori coinvolti e con le finalità proprie della tutela autoriale.

<sup>40</sup> Imprescindibile il rinvio alle riflessioni – decisive e illuminanti – di A. TURING, *Macchine calcolatrici e intelligenza*, trad. it., Torino, Einaudi, 2025.



## CAPITOLO II

### ONTOLOGIA GIURIDICA DEL DIRITTO D'AUTORE

SOMMARIO: 1. Origini gius-filosofiche della proprietà intellettuale. Riprendendo le teorie di Locke, Hegel e Bentham. – 2. L'autore come soggetto giuridico privilegiato. – 3. I confini filosofici tra opera protetta e dominio pubblico.

1. *Origini gius-filosofiche della proprietà intellettuale. Riprendendo le teorie di Locke, Hegel e Bentham*

S'è in parte già detto, il diritto d'autore – quale istituto giuridico storicamente determinato – non può essere adeguatamente compreso senza procedere ad un'analisi delle sue radici filosofiche. Diversamente da quanto si potrebbe pensare, infatti, la tutela giuridica delle opere dell'ingegno non è una conseguenza “naturale” e di per sé necessaria di un'avvenuta creazione/ideazione, è piuttosto il prodotto di precise scelte normative fondate su specifiche visioni della proprietà, della giustizia e, ovviamente, anche dell'utilità sociale. Visioni la cui genesi si colloca tra il XVII e il XIX secolo entro il contesto della modernità borghese e capitalistica e che hanno fornito un corredo di giustificazioni teorico-filosofiche a sostegno della necessità di procedere all'attribuzione in capo all'autore di alcuni diritti esclusivi sull'opera, distinguendo, così, la proprietà intellettuale dalla proprietà materiale e legittimando forme di monopolio temporaneo sulla conoscenza e sulla cultura.

Più in particolare, l'indagine genealogica sulle origini gius-filosofiche della proprietà intellettuale rivela che dietro l'apparente unitarietà del diritto d'autore si celano elaborazioni teoriche profondamente diverse. Dalla teoria lockiana del lavoro (che concepisce il diritto d'autore come ricompensa per lo sforzo creativo) a quella hegeliana della personalità (che lo interpreta come riconoscimento dell'espressione individuale), sino, ancora, alla teoria utilitaristica benthamiana (che giustifica il diritto d'autore sulla base della sua funzione di incentivo alla produzione culturale). Ciascuna di queste macro-prospettive ha influenzato l'evoluzione storica dei sistemi giuridici nazionali e continua, tuttora, a suggestionare i dibattiti contemporanei in ordine alla regolamentazione della proprietà intellettuale.

Vien da sé che nel momento in cui ci si interroga circa la considerazione da riservare alle opere generate dall'IA (se meritino protezione, a chi debba essere attribuita la titolarità dei diritti) si vanno – più o meno

consapevolmente e più o meno larvatamente – a richiamare i capitali concettuali di una (o di più di una) di queste teorie. In ipotesi, se si muove dall'adesione alla prospettiva secondo la quale il diritto d'autore premierebbe il lavoro, diviene imprescindibile chiedersi se l'attività di chi utilizza l'IA costituisca o meno un lavoro in senso proprio. Se, invece, si ritiene che il diritto d'autore vada a tutelare la personalità, bisogna stabilire se (e se sì in che misura) l'opera generata dall'IA possa manifestare ed esprimere la personalità di qualcuno. Da ultimo, se si assume che il diritto d'autore serve a incentivare la produzione culturale, allora occorre valutare se la protezione delle opere generate dall'IA favorisca (o ostacoli) questo stesso obiettivo.

Di qui, l'utilità di ricostruire criticamente le tre principali teorie gius-filosofiche della proprietà intellettuale, mettendone in luce tanto gli aspetti di continuità quanto le tensioni irrisolte. Teorie che, peraltro, l'IA mette profondamente in crisi.

Va detto subito che la teoria lockiana della proprietà – elaborata nel *Secondo Trattato sul Governo* (1690) – rappresenta uno dei fondamenti filosofici più influenti del diritto di proprietà in generale e della proprietà intellettuale in particolare. Secondo John Locke, lo stato di natura è caratterizzato da una condizione di libertà e di uguaglianza in cui tutti gli esseri umani hanno un uguale diritto rispetto alle risorse naturali che Dio ha elargito all'umanità. Detto ciò, è importante sottolineare che tale comunanza originaria, però, non implica che nessuno possa appropriarsi legittimamente di alcunché. All'opposto, ciascun individuo vanta una proprietà naturale sulla propria persona e, di conseguenza, sul lavoro del proprio corpo e sull'opera delle proprie mani<sup>1</sup>.

Questa teoria – originariamente elaborata per giustificare la proprietà fondiaria e dei beni materiali – è stata estesa alla proprietà intellettuale per mezzo del ricorso a un'analogia: così come il contadino che dissoda un terreno incolto lo rende "proprio" mescolandovi il proprio lavoro, allo stesso modo l'autore che crea un'opera letteraria o artistica acquista su di essa un diritto di proprietà in virtù di quello sforzo intellettuale che egli ha profuso. Se così, è evidente che l'opera dell'ingegno altro non sarebbe se non il frutto del lavoro mentale dell'autore e il diritto d'autore costituirebbe il precipitato logico e il doveroso riconoscimento giuridico di tale appropriazione<sup>2</sup>. L'esten-

<sup>1</sup> J. LOCKE, *Due trattati sul governo*, Torino, UTET, 2010.

<sup>2</sup> Come ha osservato Justin Hughes nel suo influente saggio sulla teoria lockiana applicata alla proprietà intellettuale, "se accettiamo che il lavoro fisico può generare diritti di proprietà sui beni tangibili, non vi è ragione di principio per negare che il lavoro intel-

sione della teoria lockiana alla tutela della proprietà intellettuale solleva una serie di problemi teorici imprevisti: *i*) intanto, mentre nel caso della proprietà fondiaria il lavoro consiste in attività fisiche osservabili e misurabili (dissodare, coltivare, edificare), nel caso della creazione intellettuale il lavoro non può essere misurato-quantificato oggettivamente (due autori possono impiegare tempi diversissimi per produrre opere di valore comparabile, un'intuizione creativa può richiedere anni di gestazione o manifestarsi in un lampo improvviso); *ii*) un secondo problema concerne la c.d. clausola lockiana della sufficienza, in ragione della quale l'appropriazione può ritenersi legittima solo se lascia agli altri risorse sufficienti e di qualità altrettanto buona (ma, nel caso delle idee e delle opere intellettuali, ogni nuova creazione, di fatto, riduce lo spazio concettuale disponibile per i creatori futuri)<sup>3</sup>; *iii*) una ulteriore criticità riguarda la distinzione tra scoperta e invenzione, storicamente, infatti, il sistema dei diritti intellettuali opera una distinzione fondamentale fra: *a*) il *copyright* (che protegge l'espressione, ovvero la forma specifica e creativa data a un'idea); *b*) il brevetto (che tutela l'invenzione, quale applicazione tecnica e funzionale di un'intuizione); *c*) la scoperta pura (che riguarda le leggi di natura, le formule matematiche o i principi fisici e che rimane solitamente esclusa da entrambe le forme di tutela, poiché si ritiene che tali verità siano universali e preesistenti all'intervento umano)<sup>4</sup>. In modo particolare, la ricostruzione lockiana sembra vacillare di fronte alla necessità di spiegare perché il semplice fatto di essere stati i primi a formulare esplicitamente una verità preesistente dovrebbe conferire un diritto di proprietà su tale acquisizione. Analogamente, anche con riguardo al dominio delle opere artistiche emerge una certa difficoltà nel sostenere che l'autore crei veramente *ex nihilo*, giacché l'evidenza storica e antropologica mostra come ogni creatore attinga necessariamente a un patrimonio culturale comune, ricombinando elementi preesistenti attraverso processi che si configurano più come trasformazione e ricontestualizzazione che come creazione assoluta. Nonostante queste criti-

lettuale possa generare diritti analoghi sulle creazioni immateriali" (J. HUGHES, "The Philosophy of Intellectual Property", in *Georgetown Law Journal*, 77(2), 1988, pp. 287-366.

<sup>3</sup> Come ha argomentato ROBERT NOZICK: "se brevetto una cura per una malattia, impedisco ad altri di copiare tale cura, e questo peggiora la loro situazione" (*Anarchy, State, and Utopia*, New York, Basic Books, 1974).

<sup>4</sup> Va detto che il punto critico della teoria di Locke emerge proprio in questo scarto: se la matematica e la scienza sono considerate scoperte di verità oggettive, l'atto del ricercatore non è una trasformazione creativa, ma un semplice rinvenimento. In questo senso, il modello lockiano – costruito sull'idea di una risorsa grezza che muta natura grazie allo sforzo individuale – fatica a giustificare l'appropriazione privata di conoscenze che non sono state "create", ma solo portate alla luce, rischiando di confondere l'estrazione di una verità con la sua effettiva produzione.

cià, tuttavia, è innegabile che la teoria lockiana abbia esercitato – e continui ad esercitare – un’influenza profonda e pervasiva sul diritto d’autore, soprattutto nei sistemi di *common law*, nei quali il legame tra lavoro e proprietà rappresenta un principio fondamentale dell’intero ordinamento giuridico<sup>5</sup>.

In questa prospettiva, la sfida posta dai sistemi di IA generativa alla teoria lockiana appare particolarmente evidente nel momento in cui si prende atto del fatto che un’opera, adesso, può essere prodotta in pochissimi secondi, senza sforzi paragonabili a quelli che erano necessari all’autore tradizionale e mediante il ricorso a processi algoritmici automatizzati che operano secondo logiche completamente diverse da quelle della creatività umana. Da par suo (e non potrebbe essere altrimenti) la teoria lockiana non offre strumenti idonei a essere utilizzati (o adattati) per determinare a chi dovrebbero spettare eventuali diritti sull’opera generata dall’intelligenza artificiale, non foss’altro perché – nel caso di specie – ci troviamo dinanzi a una frammentazione del contributo creativo che rende inapplicabile la teoria stessa, a meno che di procedere a delle forzature. Certo, il programmatore che ha creato il sistema ha certamente investito una dose di lavoro intellettuale, ma tale investimento-impegno è stato indirizzato alla progettazione dell’algoritmo generale più che alla produzione della singola opera. Lo stesso può dirsi con riguardo all’utente che ha fornito l’*input*, compiendo un’attività che può variare enormemente, da una semplice richiesta testuale di poche parole a un’elaborata ingegneria del *prompt* che richiede competenze tecniche e una certa dose di creatività. E, ovviamente, non è tutto, perché bisogna considerare anche l’impresa proprietaria dell’algoritmo che ha investito risorse economiche considerevoli nell’addestramento del modello e nella costruzione dell’infrastruttura computazionale (benché potrebbe non aver contribuito direttamente alla progettazione tecnica né alla generazione dell’*output* specifico). La frammentazione e la distribuzione del lavoro creativo tra molteplici attori rendono ardua l’identificazione di un unico soggetto a cui attribuire la proprietà (almeno secondo la logica lockiana dell’appropriazione attraverso il lavoro personale).

Ora, come si evince sin da queste primissime battute, il giurista si trova a fronteggiare istanze e tensioni che non ammettono risposte “semplici” e che non possono essere risolte mediante un’estensione analogica dei principi tradizionali. Si tratta, invero, di istanze e di tensioni che – sca-

<sup>5</sup> L’idea che l’autore meriti una ricompensa per il proprio lavoro creativo e che negargliela costituirebbe un’ingiustizia che viola principi elementari di equità, è profondamente radicata nell’intuizione morale comune e fornisce una giustificazione apparentemente solida e intuitiva per la tutela giuridica delle opere dell’ingegno.

turendo dall'implementazione di inedite realtà tecnologiche che hanno trasformato radicalmente i processi di produzione culturale – richiedono un ripensamento profondo della teoria lockiana ma anche l'individuazione di fondamenti normativi diversi e più adeguati ad uno scenario caledoscopico e altamente complesso.

Come s'è accennato, la teoria hegeliana della proprietà, sviluppata nei *Lineamenti di filosofia del diritto*, offre una giustificazione della proprietà intellettuale radicalmente diversa da quella lockiana. Hegel, infatti, guarda alla proprietà come momento necessario al processo di realizzazione della libertà individuale. Per il filosofo la persona è manifestazione di volontà libera e astratta, pura potenzialità di autodeterminazione; ma, per diventare effettivamente libera, la volontà deve oggettivarsi, deve, cioè, declinarsi – e per così dire “inverarsi” – in una realtà esterna che ne testimoni l'esistenza. Secondo questa lettura, la proprietà costituisce la prima e la più elementare forma di tale oggettivazione: appropriandosi di una cosa, la persona imprime in essa il segno della propria volontà e si riconosce come soggetto libero<sup>6</sup>.

Nello specifico caso della proprietà intellettuale, questa dinamica assume una forma peculiare e più intensa. La ragione è presto detta e, a suo modo, è intuitiva: mentre la proprietà di un oggetto materiale è “accidentale” rispetto alla personalità del soggetto-proprietario (si può alienare un terreno o un mobile senza che ciò intacchi l'identità personale), all'opposto, l'opera dell'ingegno costituisce un'esteriorizzazione diretta dello spirito del soggetto che ne è l'autore. Hegel, sul punto, è chiarissimo: “le produzioni spirituali possiedono nel loro primo formarsi esterno la peculiarità di essere, secondo la loro determinità esterna, immediatamente una cosa”<sup>7</sup>. Se così, l'opera letteraria, artistica o scientifica non è semplicemente un “prodotto” del lavoro, ma è la manifestazione esterna di facoltà intellettuali, di sentimenti e di visioni del mondo, che connotano e costituiscono la personalità stessa dell'autore. Ed è per questo preciso motivo che negare all'autore il diritto sull'opera equivale a negare il riconoscimento della sua personalità.

La teoria hegeliana fornisce il fondamento filosofico per quella che, nel diritto continentale, viene chiamata concezione personalistica del diritto d'autore, in base alla quale l'opera è inscindibilmente legata alla persona del suo creatore e deve godere di una protezione che trascende gli aspetti meramente economici. Ed è proprio a partire da questa impostazione che trovano giustificazione i c.d. diritti morali d'autore – vale a

<sup>6</sup> G.W.F. HEGEL, *Lineamenti di filosofia del diritto*, Roma-Bari, Laterza, 1996.

<sup>7</sup> *Ibidem*.

dire, il diritto di paternità, il diritto all'integrità dell'opera, il diritto di divulgazione – che caratterizzano i sistemi giuridici di *civil law* e che sono stati recepiti, almeno parzialmente, anche nella Convenzione di Berna<sup>8</sup>.

Anche la teoria hegeliana presenta alcune difficoltà applicative nel momento in cui si passa dalla giustificazione eminentemente astratta generale alla regolamentazione concreta sotto il profilo giuridico. Un primo problema riguarda la delimitazione dell'ambito di protezione: se ogni manifestazione della personalità merita tutela giuridica, allora qualsiasi forma di espressione dovrebbe essere protetta, in quanto oggettivazione della volontà. Ciò, tuttavia, contrasterebbe con il requisito dell'originalità che tutte le legislazioni pongono come prerequisito e condizione stessa della protezione. Di qui, la necessità di distinguere le opere che esprimono genuinamente la personalità dell'autore da quelle che sono mere ripetizioni di formule convenzionali o combinazioni meccaniche di elementi preesistenti. E, su questo versante, la teoria hegeliana non appare risolutiva.

A ciò si aggiunge il problema dell'alienabilità dei diritti patrimoniali. Se l'opera creativa è manifestazione della personalità, e se la personalità è inalienabile, come è possibile che l'autore possa cedere i diritti economici di sfruttamento dell'opera? Hegel stesso era consapevole di questa tensione e cercò di risolverla distinguendo tra “sostanza spirituale” dell'opera – che rimane legata indissolubilmente all'autore – e “uso esterno” dell'opera, che, invece, può essere alienato. Un distinguo, questo, che, tuttavia, si rivela assai arduo da applicare nella pratica: dove finisce la sostanza spirituale e dove inizia l'uso esterno? è coerente affermare che l'opera è manifestazione essenziale della personalità e, al contempo, ammettere che possa essere sfruttata economicamente da altri?

Un terzo problema riguarda le opere collettive o anonime. La teoria hegeliana presuppone un legame individuale tra autore e opera, ma molte creazioni sono il risultato di collaborazioni o di processi collettivi che rendono impossibile identificare un singolo creatore. Nel caso delle opere cinematografiche, ad esempio, contribuiscono il regista, lo sceneggiatore, il direttore della fotografia, il montatore, gli attori, i compositori, ciascuno con la propria visione artistica. Da qui, il dubbio: di chi è manifestazione l'opera finale? Mentre, nel caso delle opere anonime o pseudonime, può ancora parlarsi di oggettivazione della personalità quando non si conosce l'identità dell'autore?

<sup>8</sup> Come ha osservato CYRILL RIGAMONTI: “la teoria della personalità di Hegel offre la giustificazione più coerente per l'esistenza dei diritti morali, poiché considera l'opera non come un semplice bene economico ma come proiezione della personalità dell'autore” (C.P. RIGAMONTI, “Deconstructing Moral Rights”, in *Harvard International Law Journal*, 47(2), 2006, pp. 353-412).

Va detto che – se viene analizzata alla luce della teoria hegeliana – l'IA pone una sfida ancora più radicale. Daltronde, se l'opera è un'oggettivazione della personalità – e se l'IA non ha personalità nel senso hegeliano (non è spirito libero che si autodetermina) – le sue produzioni non dovrebbero ricevere alcuna tutela. Eppure, questa conclusione si scontra con l'evidenza che le opere generate dall'IA possono essere esteticamente apprezzabili, socialmente utili ed economicamente preziose. Inoltre, si potrebbe argomentare che, anche quando l'opera è generata da un algoritmo, vi è comunque un elemento di personalità umana che si manifesta: quella del programmatore che ha progettato il sistema, oppure quella dell'utilizzatore che ha selezionato i parametri e guidato il processo generativo. Se si accetta questo argomento, si ridimensiona notevolmente la teoria hegeliana. Ragion per cui la personalità non si manifesterebbe più direttamente nell'opera, ma solo indirettamente attraverso la mediazione tecnologica.

In breve: la teoria hegeliana offre una giustificazione filosoficamente profonda della proprietà intellettuale, radicandola nella struttura stessa della libertà umana e della personalità. In tal modo, essa spiega perché la tutela dell'opera non possa essere meramente economica, ma debba includere il riconoscimento di quel particolare legame spirituale che si genera tra l'autore e la sua creazione. Tuttavia – così come avviene nel caso della teoria lockiana – l'avvento dell'IA fa barcollare la ricostruzione, in quanto fa affiorare tutti i limiti di una concezione che presuppone la presenza di un soggetto umano dotato di personalità.

Di qui, la necessità di interrogarsi sulla possibilità di preservare gli aspetti più preziosi della teoria hegeliana – in particolare l'attenzione alla dimensione non patrimoniale della creazione – anche in un contesto in cui il creatore non è più esclusivamente umano.

La terza grande teoria gius-filosofica della proprietà intellettuale è quella utilitaristica, elaborata principalmente da Jeremy Bentham e adottata come fondamento nei sistemi di *common law*, in particolare negli Stati Uniti. L'utilitarismo radica la proprietà intellettuale sulla sua utilità sociale. Secondo Bentham, il fine ultimo di ogni istituzione giuridica deve essere la massimizzazione della felicità complessiva, intesa come saldo positivo tra piaceri e dolori della collettività. Il diritto, in questa prospettiva, non serve a realizzare la giustizia in senso astratto, ma a promuovere il benessere concreto degli individui<sup>9</sup>. Ad essere più precisi, Bentham ritiene che la proprietà sia un'aspettativa di utilità, che non esiste in na-

<sup>9</sup> J. BENTHAM, *An Introduction to the Principles of Morals and Legislation*, Oxford, Clarendon Press, 1789.

tura, in quanto è un prodotto della legge. Se applicato alla proprietà intellettuale, il principio utilitaristico conduce a una giustificazione di tipo strumentale: il diritto d'autore è legittimo non perché l'autore abbia un diritto naturale sull'opera o perché essa manifesti la sua personalità, ma perché l'attribuzione di diritti esclusivi incentiva la produzione di opere che altrimenti non verrebbero create, incrementando così la disponibilità complessiva di beni culturali e il benessere sociale<sup>10</sup>.

La Costituzione degli Stati Uniti recepisce esplicitamente questa logica quando attribuisce al Congresso il potere di promuovere il progresso della scienza e delle arti utili, garantendo per tempi limitati agli autori e agli inventori il diritto esclusivo ai loro rispettivi scritti e scoperte (Art. I, Sez. 8, Cl. 8). La finalità della protezione non è ricompensare l'autore o riconoscere la sua dignità, ma incentivare la creazione di opere che arricchiscano il patrimonio culturale della nazione. In quest'ottica il diritto d'autore costituisce un male necessario – un monopolio temporaneo che limita l'accesso alla conoscenza – giustificato unicamente dalla considerazione che senza tale incentivo la produzione culturale sarebbe insufficiente.

La teoria utilitaristica presenta alcuni vantaggi rispetto alle teorie lockiana e hegeliana. In primo luogo, fornisce un criterio operativo per valutare le politiche di proprietà intellettuale: occorre chiedersi se un dato regime normativo massimizzi effettivamente la produzione culturale e il benessere sociale. In secondo luogo, riconosce esplicitamente che il diritto d'autore comporta costi sociali (limitazione dell'accesso alle opere, ostacoli alla creatività derivativa, aumento dei prezzi) che devono essere bilanciati con i benefici (incentivo alla creazione). In terzo luogo, giustifica naturalmente la limitazione temporale della protezione: poiché il diritto esclusivo è solo uno strumento per incentivare la creazione, esso deve durare solo il tempo necessario a tale scopo, dopodiché l'opera deve entrare nel dominio pubblico per massimizzare l'utilità sociale.

Tuttavia, anche la teoria utilitaristica non è esente da limiti strutturali che ne condizionano la tenuta sul piano normativo. Un primo e fondamentale problema attiene alla verificabilità empirica del suo presupposto centrale: l'assunto secondo il quale il *copyright* massimizzi il benessere sociale complessivo è assunto come dato, ma non è mai stato dimostrato

<sup>10</sup> Come osservato da WILLIAM LANDES e RICHARD POSNER nella loro analisi economica del copyright, “la giustificazione economica fondamentale del diritto d'autore è che, in assenza di protezione giuridica, i creatori di opere espressive sarebbero incapaci di recuperare i loro investimenti in creazione e produzione, poiché chiunque potrebbe copiare gratuitamente l'opera una volta pubblicata” (W.M. LANDES, R.A. POSNER, *The Economic Structure of Intellectual Property Law*, Cambridge, MA, Harvard University Press, 2003).

in modo conclusivo dalla ricerca empirica<sup>11</sup>. Gli studi dedicati all'impatto delle diverse variabili del sistema – durata della protezione, ampiezza dei diritti esclusivi, articolazione delle eccezioni e limitazioni – sulla produzione culturale e sull'innovazione pervengono a conclusioni sistematicamente divergenti, riflettendo la difficoltà metodologica di isolare il contributo del *copyright* dagli altri fattori economici, sociali e tecnologici che concorrono a determinare i livelli di creatività e di accesso alla cultura. In questa condizione di incertezza empirica persistente, il criterio dell'utilità sociale rischia di rivelarsi uno strumento duttile e, in quanto tale, suscettibile di essere piegato retoricamente a sostegno delle posizioni più diverse – ora a favore di una protezione più intensa, ora a sostegno di una sua drastica riduzione – senza che l'analisi empirica disponibile sia in grado di arbitrare in modo cogente tra le opzioni in campo. Il paradosso della teoria utilitaristica applicata al diritto d'autore è, dunque, che il suo stesso rigore metodologico – la pretesa di fondare le scelte normative su dati misurabili e su previsioni verificabili – finisce per rivoltarsi contro di essa nel momento in cui i dati sono controversi e le previsioni incerte.

Sotto un diverso profilo, emerge la questione del nesso tra il cosiddetto incentivo *ex ante* e l'accesso *ex post*. Nel dettaglio, la teoria utilitaristica ritiene che sia necessario offrire agli autori un incentivo economico per indurli a creare, ma questa assunzione è empiricamente dubbia fragile e sconfessabile. Molti autori, ad esempio, creano per motivazioni non economiche – quali: desiderio di espressione, ricerca di riconoscimento, passione artistica – e continuerebbero a farlo anche in totale assenza di protezione giuridica<sup>12</sup> (e di conseguente riconoscimento economico). Se questo è vero, allora i costi del diritto d'autore potrebbero anche supera-

<sup>11</sup> L'inconcludenza degli studi empirici sull'impatto del diritto d'autore sulla produzione culturale è documentata da un'ampia letteratura. Per una panoramica critica si vedano, in ambito italiano: P. AUTERI, G. FLORIDIA, V. MANGINI, G. OLIVIERI, M. RICOLFI, P. SPADA, *Diritto industriale. Proprietà intellettuale e concorrenza*, Torino, Giappichelli, 2016; M. RICOLFI, *Il diritto d'autore*, in N. ABRIANI, G. COTTINO, M. RICOLFI, *Diritto industriale*, Padova, CEDAM, 2001. In ambito internazionale: R.M. HURT, R.M. SCHUCHMAN, "The Economic Rationale of Copyright", in *American Economic Review*, 56(1/2), 1966, pp. 421-432; S. BREYER, "The Uneasy Case for Copyright: A Study of Copyright in Books, Photocopies, and Computer Programs", in *Harvard Law Review*, 84(2), 1970, pp. 281-351; I. PNG, "Copyright: A Plea for Empirical Research", in *Review of Economic Research on Copyright Issues*, 3(2), 2006, pp. 3-13; nonché la rassegna critica di R. TOWSE (a cura di), *Copyright in the Cultural Industries*, Cheltenham, Edward Elgar, 2002.

<sup>12</sup> Come ha osservato Benkler nel suo studio sulla produzione *peer-to-peer*, "una parte significativa della creazione culturale contemporanea avviene al di fuori della logica del mercato, attraverso forme di collaborazione non proprietaria" (Y. BENKLER, *The Wealth*

re i suoi benefici, poiché si limiterebbe l'accesso alle opere senza che ciò comporti un effettivo incremento della produzione.

Un terzo problema riguarda la distribuzione degli effetti del diritto d'autore. Invero, anche ammettendo che il *copyright* massimizzi l'utilità aggregata, occorre chiedersi come tale utilità sia distribuita tra i diversi gruppi sociali. Se la protezione giuridica avvantaggia principalmente le grandi *corporation* dell'industria culturale a scapito degli autori individuali e del pubblico, può ancora sostenersi che essa realizzi il benessere collettivo?

L'IA, inoltre, pone alla teoria utilitaristica interrogativi specifici. Se l'obiettivo del diritto d'autore è incentivare la produzione culturale, occorre chiedersi se la protezione delle opere generate dall'IA serva effettivamente tale scopo. Da un lato, si potrebbe argomentare che attribuire diritti sulle opere generate dall'IA incentiverebbe gli investimenti in ricerca e sviluppo di sistemi sempre più sofisticati, aumentando, così, la disponibilità di contenuti culturali. Dall'altro lato, si potrebbe sostenere che la protezione di tali opere ridurrebbe lo spazio creativo disponibile per gli autori umani, generando un effetto netto negativo sulla produzione culturale complessiva. La teoria utilitaristica non fornisce una risposta *a priori* a questa questione, ma richiede un'analisi empirica degli effetti concreti delle diverse opzioni normative.

Un quarto ordine di difficoltà è segnalato dall'analisi economica del diritto. Applicando alla proprietà intellettuale il concetto di "tragedia degli anticommons" elaborato da Michael Heller<sup>13</sup>, si può sostenere che estendere diritti esclusivi a ogni *output* algoritmico potrebbe produrre l'effetto opposto a quello auspicato dalla teoria utilitaristica: anziché incentivare la produzione culturale, la frammentazione di un numero potenzialmente illimitato di diritti esclusivi su opere generate a costo marginale tendente a zero genererebbe costi di transazione – nel senso

*of Networks: How Social Production Transforms Markets and Freedom*, New Haven, Yale University Press, 2006).

<sup>13</sup> Il concetto di "tragedia degli anticommons" è stato elaborato da M.A. HELLER, "The Tragedy of the Anticommons: Property in the Transition from Marx to Markets", in *Harvard Law Review*, 111(3), 1998, pp. 621-688, e applicato alla proprietà intellettuale in M.A. HELLER, R.S. EISENBERG, "Can Patents Deter Innovation? The Anticommons in Biomedical Research", in *Science*, 280(5364), 1998, pp. 698-701. La nozione descrive la situazione in cui una molteplicità di diritti esclusivi frammentati su una medesima risorsa determina una sottoutilizzazione della stessa, poiché i costi necessari per ottenere tutte le autorizzazioni richieste eccedono il valore atteso dall'utilizzo. Sul concetto di costi di transazione quale strumento di analisi del diritto della proprietà, il riferimento fondamentale rimane R.H. COASE, "The Problem of Social Cost", in *Journal of Law and Economics*, 3, 1960, pp. 1-44.

coasiano<sup>14</sup> del termine – talmente elevati da rendere la gestione delle licenze economicamente insostenibile, ostacolando di fatto la circolazione della conoscenza. In altri termini, la logica utilitaristica che giustifica il monopolio temporaneo come incentivo alla creazione rischia di auto-confutarsi nel momento in cui viene estesa a produzioni algoritmiche seriali, per le quali il problema non è la scarsa disponibilità di contenuti, ma semmai la loro sovrabbondanza. Questo rilievo critico non conduce, di per sé, a soluzioni normative univoche: stabilire se e in quale misura le opere generate dall'IA debbano ricevere tutela è questione che richiede un bilanciamento tra valori – incentivo all'innovazione, accesso alla cultura, giustizia distributiva – che sarà affrontato nei capitoli successivi. Ciò che preme sottolineare in questa sede è che la teoria utilitaristica, lungi dal fornire una risposta univoca, espone essa stessa le proprie aporie non appena si confronta con le specificità della creatività algoritmica.

## 2. *L'autore come soggetto giuridico privilegiato*

Le teorie gius-filosofiche qui molto sinteticamente richiamate convergono – pur da prospettive diverse – nel riconoscere all'autore uno status giuridico privilegiato. Che si fondi tale privilegio sul lavoro creativo, sulla manifestazione della personalità oppure sulla funzione di incentivo, il risultato è comunque sia l'attribuzione in capo ad un soggetto determinato di diritti esclusivi su beni immateriali che, per loro natura, potrebbero essere utilizzati simultaneamente da un numero indefinito di individui pur senza consumarsi. Questa attribuzione non è affatto ovvia o naturale: richiede una costruzione normativa che identifichi chi sia l'autore, quali facoltà gli spettino e per quanto queste stesse facoltà debbano essere protette. Il diritto d'autore, in altre parole, non si limita a riconoscere una situazione pre-esistente, ma costituisce attivamente la figura dell'autore come soggetto giuridico dotato di prerogative speciali.

Tale costruzione normativa presuppone, a sua volta, determinati assunti antropologici e ontologici circa la natura della creazione, quella del creatore e, in ordine, al loro rapporto. Il diritto d'autore moderno è edificato sull'idea che il creatore sia necessariamente una persona umana, dotata di capacità cognitive, di volontà autonoma e di sensibilità estetica. Questi attributi, considerati essenziali per la produzione di opere originali, giustificano il nesso tra personalità giuridica e capacità creativa che

<sup>14</sup> Vale a dire, nel segno delle teorie elaborate dal premio Nobel RONALD COASE.

costituisce il fondamento del sistema autoriale<sup>15</sup>. Tale visione dell'autorialità romantica testimonia che il diritto d'autore moderno è figlio del Romanticismo (l'autore come genio solitario) ecco perché si potrebbe sostenere che l'IA non sia solo una minaccia tecnica, ma la fine dell'estetica romantica nel diritto d'autore.

L'avvento dell'IA mette in discussione il presupposto antropocentrico. Se un sistema artificiale può generare opere indistinguibili da quelle umane quanto a qualità estetica e originalità formale, il nesso tra personalità umana e capacità creativa si rivela contingente anziché necessario.

Da qui, la necessità e l'urgenza per il giurista – e in modo particolare per il filosofo del diritto – di esaminare criticamente l'idea e gli stessi fondamenti normativi dell'autorialità, individuando e valorizzando quegli elementi che sono essenziali e quelli che, invece, pur essendo storicamente determinati, risultano potenzialmente modificabili. Il tutto, non già nell'intento di minimizzare o pretermettere la figura dell'autore, disconoscendone l'importanza, quanto, piuttosto, al fine di ripensarla e di adeguarla alle sfide tecnologiche. In questa prospettiva, occorre analizzare come il diritto abbia storicamente costruito la figura dell'autore attraverso norme che definiscono chi possa essere considerato creatore e quali requisiti debba soddisfare l'opera per essere protetta. Anzitutto, ci si interrogherà sul nesso tra personalità giuridica e capacità creativa, ragionando sulla possibilità di ipotizzare una tutela che prescindano dal legame diretto con una persona; in secondo luogo, ci si soffermerà sulla centralità antropologica nel sistema del *copyright*.

Come si è detto, la figura dell'autore è il prodotto di una lunga evoluzione storica, connessa all'idea che il creatore di un'opera possa/debba vantare su di essa dei diritti individuali riconosciuti e tutelati dall'ordinamento giuridico. Nelle società antiche e medievali, tale idea era pressoché assente: l'opera circolava anonimamente, veniva liberamente copiata e modificata e l'eventuale fama del creatore dipendeva dal riconoscimento sociale anziché da un diritto giuridicamente azionabile. Il periodo storico che segna la nascita del diritto d'autore è quello compreso tra il XVII e il XVIII secolo e coincide con l'affermazione di una concezione individualistica della creazione, in base alla quale l'autore è il titolare originario di una serie di prerogative fondamentali ed esclusive.

<sup>15</sup> Come ha osservato Jaszi, "il diritto d'autore non tutela semplicemente opere, ma autori: soggetti individuali la cui identità è indissolubilmente legata alle loro creazioni" (P. JASZI, "Toward a Theory of Copyright: The Metamorphoses of 'Authorship'", in *Duke Law Journal*, 1991(2), 1991, pp. 455-502).

Il primo passo in questa costruzione normativa fu il riconoscimento dell'autore come soggetto distinto dagli editori e dai librai, che controllavano la produzione e la distribuzione delle opere. Lo *Statuto di Anna* (1710) – considerato la prima forma di *copyright* moderno – attribuiva agli autori (e non più alle corporazioni di stampatori) il diritto esclusivo di stampare e di vendere le proprie opere per un periodo di quattordici anni, che era rinnovabile una volta. Questa innovazione normativa presupponeva che esistesse un soggetto identificabile – per l'appunto identificato come l'autore – al quale potesse essere attribuita la paternità dell'opera e, di conseguenza, la titolarità di taluni diritti<sup>16</sup>.

La legislazione rivoluzionaria francese<sup>17</sup> rafforzò ulteriormente la centralità riconosciuta alla figura dell'autore, soprattutto attraverso le leggi del 1791 e del 1793, che segnarono una tappa decisiva nella costruzione del *droit d'auteur*. Il *Code civil* napoleonico del 1804, pur non dedicando ampio spazio alla proprietà letteraria, consolidò sul piano generale questa impostazione individualistica, che sarebbe poi stata recepita nei codici di molti altri paesi e avrebbe favorito l'ulteriore elaborazione della tutela autorale.

Nel corso del XIX secolo, la costruzione normativa dell'autorialità si arricchì di ulteriori elementi. La Convenzione di Berna del 1886 – che per la prima volta istituì una protezione internazionale del diritto d'autore – sancì, da un lato, il principio del trattamento nazionale, in forza del quale gli autori godono negli altri paesi dell'Unione dei diritti riconosciuti ai cittadini di quegli stessi paesi e dei diritti specialmente previsti dalla Convenzione; dall'altro lato, il principio secondo cui il godimento e l'esercizio di tali diritti non sono soggetti ad alcuna formalità. Ciò significava che l'autore si vedeva riconosciuto come tale in virtù del mero atto creativo, indipendentemente da registrazioni o da altre procedure amministrative. Inoltre, la *Convenzione* sancì il diritto morale dell'autore di rivendicare la paternità dell'opera e di opporsi a modifiche che ledessero

<sup>16</sup> Come ha notato MARK ROSE nel suo studio sulla costruzione della figura dell'autore, “*lo Statuto di Anna non si limitò a creare un nuovo diritto, ma contribuì a forgiare una nuova identità sociale: quella dell'autore professionista che vive del proprio lavoro creativo*” (*Authors and Owners: The Invention of Copyright*, Cambridge (MA), Harvard University Press, 1993).

<sup>17</sup> I decreti del 1791 e del 1793 proclamarono solennemente che “*la più sacra, la più personale di tutte le proprietà è l'opera frutto del pensiero di uno scrittore*”, introducendo così nel diritto positivo la concezione personalistica della proprietà intellettuale. Significativamente, questa retorica dell'autore-proprietario emergeva in un contesto rivoluzionario che aveva abolito i privilegi feudali e le corporazioni: l'autore veniva costruito come individuo borghese che, attraverso il proprio lavoro intellettuale, acquisiva un diritto naturale equiparato alla proprietà materiale.

il proprio onore o la propria reputazione, riconoscendo che il legame tra l'autore e la sua opera trascende la dimensione meramente economica.

È importante osservare che questa costruzione normativa dell'autorialità presupponeva determinati requisiti che non tutti i soggetti potevano soddisfare: *i*) l'autore doveva essere una persona fisica (le persone giuridiche, pur potendo acquisire diritti d'autore per cessione o per legge, non potevano essere considerate autrici originarie); *ii*) l'autore doveva avere capacità giuridica (i minori e gli incapaci, pur potendo creare opere, necessitavano della rappresentanza legale per esercitare i diritti); *iii*) l'autore doveva essere identificabile, almeno in linea di principio (le opere anonime o pseudonime ricevevano una tutela speciale proprio perché l'autore, pur esistendo, non era noto); *iv*) l'autore doveva necessariamente essere umano (nessuna legislazione prendeva in considerazione la possibilità che animali o entità artificiali potessero creare opere proteggibili).

Questo impianto concettuale e normativo – ancorché diffuso e consolidato – non è immune da tensioni e da problemi. Una prima questione riguarda le opere collettive o collaborative e, nel dettaglio, ha a che vedere con la difficoltà di individuare l'autore quando molteplici soggetti contribuiscono alla creazione dell'opera. In questo caso, le legislazioni risolvono il problema distinguendo tra co-autorialità (quando i contributi sono inscindibili), opere composte (quando i contributi sono distinguibili) e opere collettive (quando un soggetto coordina i contributi altrui). E già qui emerge come il modello dell'autore individuale – per quanto ideologicamente dominante – si dimostra incapace a dar conto della complessità dei processi creativi reali.

Un ulteriore nodo riguarda le opere create su commissione (o nell'ambito di rapporti di lavoro subordinato), che, in molti ordinamenti vengono ritenute appartenere al datore di lavoro o al committente, anche se l'autore (in senso creativo) rimane il dipendente. Una dissociazione tra autorialità e titolarità dei diritti, questa, che denota come il sistema giuridico tenga in notevole considerazione la praticità economica: in pratica quello che conta non è tanto riconoscere il genio creativo, ma attribuire diritti economici a chi ha sostenuto gli investimenti.

L'IA – come è *ictu oculi* evidente – sollecita delle criticità che non possono essere risolte semplicemente in via analogica, estendendo le categorie esistenti. In particolare, nel momento in cui un sistema di IA generi autonomamente un'opera senza intervento umano significativo, nessuna delle soluzioni tradizionali – la tutela riservata all'opera anonima, l'idea di opera collettiva, o, ancora, il rinvio alla nozione di opera su commissione – appare soddisfacente. Occorre, dunque, procedere ad un

ripensamento radicale di chi e/o di che cosa possa essere ritenuto quale autore e, soprattutto, appare necessario, valutare se abbia ancora un senso la centralità di questa figura (eminentemente soggettiva/personalistica/individuale) a guidare tutto l'impianto sistematico del diritto d'autore. Per forza di cose, questo ripensamento porta con sé una riconfigurazione del nesso tra l'idea di personalità giuridica e quella di capacità creativa.

In questa sede è importante ribadire che il legame tra la personalità e la creatività è talmente radicato che le legislazioni non sentono più nemmeno il bisogno di esplicitarlo: invero, è dato per scontato che l'autore di un'opera di ingegno sia una persona fisica. Nonostante ciò, un'analisi più attenta rivela che tale nesso – ancorché diffuso e inveterato – non è affatto necessario né imprescindibile, bensì contingente. La personalità giuridica, difatti, non è un attributo per così dire “naturale” ma una qualificazione normativa: l'ordinamento, cioè, decide chi riconoscere come persona e a chi attribuire la capacità giuridica. Tanti sono gli esempi che si possono rintracciare nel corso della storia<sup>18</sup>.

Ora, ai fini del nostro percorso di indagine, è evidente che nel momento in cui si muove da tale presa di coscienza, non rimane alcuna ragione – almeno in via di principio – per escludere che entità non-umane possano essere riconosciute come soggetti di diritto e possano vedersi riconosciuti titolari anche del diritto d'autore.

Venendo alla questione in esame – ovvero alla possibilità/necessità di ripensare il diritto d'autore alla luce degli scenari dischiusi dall'IA – è interessante sottolineare che alcuni giuristi sostengono che l'estensione della personalità giuridica alle IA rappresenta l'unica soluzione coerente al problema delle opere generate autonomamente dai sistemi algoritmici. Invero, se l'IA crea opere che soddisfano il requisito dell'originalità, allora essa dovrebbe comunque essere riconosciuta come autrice e titolare dei diritti, analogamente a quanto avviene per le persone giuridiche nel diritto societario. Lawrence Solum – in un pionieristico ed influente articolo del 1992 – propose di introdurre la categoria della *persona artificiale legale* (*legal artificial person*) per designare quei sistemi di IA sufficientemente sofisticati da meritare lo *status* di “soggetto” di diritto<sup>19</sup>. Una

<sup>18</sup> Storicamente, non tutti gli esseri umani sono stati considerati persone (si pensi agli schiavi nel diritto romano o agli stranieri privi di *status civitatis*) e non solo gli esseri umani sono riconosciuti come persone (le società, le fondazioni, gli enti pubblici sono persone giuridiche pur non avendo corpo né coscienza).

<sup>19</sup> Come ha osservato BAYERN, “attribuire personalità giuridica a un'entità priva di coscienza significa creare un soggetto fittizio che altri (programmatori, proprietari, utilizzatori) useranno per i propri scopi, eludendo potenzialmente responsabilità e obbli-

proposta, questa, che per altro è stata recentemente ripresa nel contesto del dibattito europeo sull'*electronic personhood* per i robot avanzati.

Non stupisce che l'attribuzione di personalità giuridica all'IA sollevi problemi teorici e pratici di non poco conto. Sul piano teorico, la personalità giuridica è tradizionalmente giustificata dalla necessità di tutelare interessi che solo gli esseri coscienti e senzienti possono avere. Le persone giuridiche (società, fondazioni) sono finzioni normative che servono a organizzare e proteggere interessi umani collettivi, ma una IA, almeno allo stato attuale della tecnologia, non ha interessi propri da tutelare<sup>20</sup>.

Spostandosi sul piano pratico, poi, riconoscere l'IA come autrice genererebbe una serie di conseguenze paradossali. Intanto sorgerebbe il problema dell'esercizio dei diritti morali: un sistema algoritmico è strutturalmente incapace di rivendicare la paternità di un'opera o di opporsi a modifiche lesive della propria reputazione, sicché tali diritti resterebbero privi di un titolare in grado di esercitarli concretamente. Sul versante della responsabilità, la qualificazione dell'IA come autrice aprirebbe il problema speculare della sua possibile qualificazione come contraffattrice nei casi di copia illegittima di opere altrui, con la conseguenza – di difficile soluzione – dell'assenza di un patrimonio su cui rivalersi nel caso in cui si renda necessario procedere con un risarcimento del danno. Già da questi primissimi rilievi, appare chiaro che l'attribuzione di personalità giuridica all'IA può rischiare di creare più problemi di quanti, invece, ne risolve.

Una possibile alternativa alla soluzione di ricorrere all'elaborazione e all'attribuzione della personalità artificiale consiste, per un verso, nel mantenere fermo il principio secondo il quale solo le persone fisiche possono essere ritenute autrici di un'opera e, simultaneamente, per un altro verso, nell'estendere il concetto di autorialità in modo da riuscire ad includere nella categoria anche coloro i quali utilizzano l'IA come strumento creativo. In pratica, seguendo questa tesi, nel momento in cui una persona ricorre a un sistema di IA per generare un'opera fornendo *input*, parametri e indicazioni creative, quella stessa persona dovrebbe essere riconosciuta come autrice, analogamente a quanto avviene per l'artista che usa un pennello o un *software* di grafica<sup>21</sup>. Si tratta, però, di una soluzione che solleva

ghi" (L.B. SOLUM, "Legal Personhood for Artificial Intelligences", in *North Carolina Law Review*, 70(4), 1992, pp. 1231-1287); cfr. anche European Parliament, Resolution of 16 February 2017 with recommendations to the Commission on *Civil Law Rules on Robotics* (2015/2103(INL)), P8\_TA(2017)0051.

<sup>20</sup> S. BAYERN, "The Implications of Modern Business-Entity Law for the Regulation of Autonomous Systems", in *Stanford Technology Law Review*, 19, 2015, pp. 93-112.

<sup>21</sup> È interessante ricordare il caso *Thaler v. Perlmutter* – deciso dal tribunale distret-

a sua volta ulteriori problemi, fra tutti, quello relativo alla determinazione del grado di intervento umano necessario affinché si possa parlare di autorialità. Se, infatti, basta fornire un *prompt* testuale a un sistema generativo, allora chiunque usi *ChatGPT* o *DALL-E* sarebbe autore delle opere prodotte. Un'interpretazione assai ampia, che sembra svilire il concetto di creatività. Del resto – ed è lapalissiano – scrivere “genera un'immagine di un tramonto” non equivale nemmeno lontanamente all'azione di dipingere un tramonto. D'altra parte, nel momento in cui si richiede un intervento umano più sostanziale – del tipo: selezione tra molteplici *output*, *editing*, integrazione con elementi creati manualmente – si rischia di escludere dalla protezione opere generate con un contributo umano minimo, ma comunque apprezzabili. Un'altra via, più radicale, consisterebbe nel mettere in discussione il nesso tra autorialità e diritto esclusivo. In questa prospettiva, quando l'IA genera opere senza che vi sia un autore umano in senso proprio, tali opere possono entrare direttamente nel dominio pubblico, liberamente utilizzabili da chiunque. Questa soluzione rispetta il principio in base al quale solo gli esseri umani possano essere autori e non richiede di forzare il concetto di autorialità per ricomprendervi situazioni spurie. Fermo restando che, per come opera, tale soluzione eliminerebbe qualsiasi incentivo economico allo sviluppo di sistemi di IA generativa, implicando possibili effetti negativi sull'innovazione tecnologica.

In breve, nessuna delle soluzioni proposte risulta pienamente soddisfacente. E resta il fatto – evidente e incontrovertibile – che il nesso tra personalità giuridica e capacità creativa si palesa del tutto inadeguato a misurarsi e a dar conto dei fenomeni di creatività artificiale. Di qui, il bisogno di individuare soluzioni in grado di bilanciare adeguatamente i diversi valori in gioco, ovvero sia: la tutela della creatività umana, l'incentivo all'innovazione, l'accesso alla cultura, la coerenza sistematica del diritto. Un bilanciamento, questo, che implica un radicale ripensamento della prospettiva eminentemente antropologica implementata dal sistema del *copyright*.

Occorrè però, osservare che l'antropocentrismo, che connota il sistema del diritto d'autore, si traduce nel fatto che ogni sua categoria, ogni suo istituto, ogni sua implicazione (o ricaduta) presuppone che la creazione sia un'attività umana e che l'autore sia una persona dotata di quelle facoltà cognitive ed emotive che costituiscono la cifra stessa dell'umano.

tuale del District of Columbia – che ha rigettato la richiesta di registrazione di un *copyright* per un'opera generata autonomamente da IA, affermando che “il copyright richiede un elemento di creatività umana che deve provenire da un essere umano” (United States District Court D.D.C., *Thaler v. Perlmutter*, No. 22-cv-1564 (18 agosto 2023)).

Ovviamente, la visione antropologica non è altro che un precipitato della concezione moderna dell'individuo come soggetto autonomo, razionale e creativo che, attraverso il proprio lavoro intellettuale, si realizza e contribuisce al progresso della società<sup>22</sup>.

Questa centralità antropologica si manifesta in molteplici aspetti della disciplina giuridica. Innanzitutto, nel requisito dell'originalità: un'opera è protetta solo se reca l'impronta della personalità dell'autore e la personalità è un attributo specificamente umano. Si noti qui che l'IA sfida il requisito dell'originalità intesa come "scelta creativa dell'autore" (visione soggettiva europea) a favore di una visione di "risultato originale" (visione più oggettiva/anglosassone). In secondo luogo, nei diritti morali: solo un essere dotato di dignità e reputazione può rivendicare la paternità di un'opera o opporsi a modifiche lesive. In terzo luogo, nella durata della protezione: il termine di tutela è parametrato alla vita dell'autore (vita dell'autore più settant'anni), presupponendo così che l'autore sia un essere mortale. In quarto luogo, nelle eccezioni e limitazioni: molte di esse (come la citazione per critica, l'uso per fini didattici, la parodia) presuppongono attività tipicamente umane.

Nel contesto del diritto d'autore, abbandonare la prospettiva antropocentrica – che, per molti aspetti, accompagna l'elaborazione normativo-ordinamentale a ragione – significherebbe riconoscere che la creatività non è più un monopolio dell'umano e che le opere di valore possono essere generate anche da sistemi artificiali. Questo passaggio, tuttavia, non implica necessariamente la necessità di attribuire personalità giuridica all'IA o di equiparare completamente le sue produzioni a quelle umane. Si tratta, piuttosto, di elaborare categorie che permettano di tutelare l'innovazione tecnologica pur senza per ciò stesso negare la peculiarità della creatività umana. Alcune proposte avanzate in dottrina prevedono l'introduzione di forme di protezione *sui generis* per le opere generate dall'IA, analoghe a quei diritti (già esistenti) che sono stati introdotti con riferimento alle banche dati o alle prestazioni degli artisti interpreti. Il punto è che se si ammette che anche le macchine possono creare, si rischia di svilire il valore della creatività umana, riducendola a mero processo computazionale. Inoltre, si rischia di favorire le grandi *corporation* tecnologiche che controllano i sistemi di IA più avanzati, a

<sup>22</sup> Come è stato osservato: "il copyright non protegge semplicemente opere, ma istituisce e riproduce una certa visione dell'umano: l'umano come creatore originale" (M. BIAGIOLI, *Nature Engaged: Science in Practice from the Renaissance to the Present*, New York, Palgrave Macmillan, 2011).

scapito degli autori individuali che vedrebbero ridotto lo spazio per la propria creatività<sup>23</sup>.

L'idea guida è che il diritto d'autore deve continuare a tutelare primariamente la creatività umana, ma può prevedere forme di protezione complementari per le opere generate con l'ausilio o attraverso l'intervento dell'IA, purché sia garantito un giusto bilanciamento tra incentivo all'innovazione e tutela degli autori umani. Questo approccio richiede di distinguere situazioni diverse: *i*) opere create da esseri umani con l'ausilio di strumenti tecnologici (che rimangono pienamente protette); *ii*) opere co-create da esseri umani e IA (che potrebbero ricevere una protezione ridotta o condizionata); *iii*) opere generate autonomamente dall'IA (che potrebbero non ricevere protezione o ricevere una forma di tutela *sui generis*).

In definitiva, la sfida posta dall'IA al diritto d'autore non è solo tecnico-giuridica, ma investe questioni filosofiche fondamentali circa la natura dell'umano, il valore della creatività e i fini del diritto. Affrontare tale sfida richiede un approccio interdisciplinare che integri analisi giuridica, riflessione filosofica e comprensione tecnologica. Solo attraverso questo sforzo collettivo sarà possibile elaborare soluzioni normative che siano al contempo eticamente fondate, tecnologicamente consapevoli e giuridicamente praticabili.

### 3. I confini filosofici tra opera protetta e dominio pubblico

Il diritto d'autore è strutturalmente caratterizzato da un'ambivalenza di fondo: da un lato attribuisce all'autore diritti esclusivi sull'opera (creando così una forma di monopolio temporaneo sulla conoscenza e sulla cultura) e, dall'altro, riconosce l'esistenza di un dominio pubblico (uno spazio di libera circolazione e utilizzazione delle idee, delle informazioni e delle espressioni non protette o non più protette). La determinazione del confine tra ciò che è riconducibile sotto l'egida diritto esclusivo (e può divenire oggetto di appropriazione) e ciò che, invece, deve rimanere nella disponibilità comune non è una questione meramente tecnica, ma investe scelte filosofiche e politiche fondamentali circa il rapporto tra l'interesse privato e l'utilità collettiva, tra incentivo alla creazione e accesso alla conoscenza, tra libertà individuale e progresso sociale. Tale confine è tracciato attraverso una serie di distinzioni concettuali che operano

<sup>23</sup> Come ha avvertito HESSE: "il pericolo di un copyright post-umano è che esso finisca per servire gli interessi delle macchine (o meglio, dei loro proprietari) piuttosto che quelli degli esseri umani" (C. HESSE, "The Rise of Intellectual Property, 700 B.C.-A.D. 2000: An Idea in the Balance", in *Daedalus*, 131(2), 2002, pp. 26-45).

su piani diversi. Innanzitutto, la distinzione tra opera originale e mera copia o imitazione: solo ciò che presenta un carattere creativo sufficiente merita protezione. In secondo luogo, la distinzione tra espressione e idea: il diritto d'autore tutela le forme espressive ma non i contenuti ideali sottostanti. In terzo luogo, la distinzione temporale: la protezione è limitata nel tempo, dopodiché l'opera entra nel dominio pubblico. Infine, la distinzione tra utilizzazioni riservate e utilizzazioni libere: anche durante il periodo di protezione, esistono eccezioni e limitazioni che permettono determinati usi senza necessità di autorizzazione.

Queste distinzioni, per quanto consolidate nella prassi giuridica, sollevano interrogativi filosofici profondi relativi alla natura stessa dell'originalità, al confine tra idea ed espressione, alla giustificazione della durata limitata della protezione e ai criteri che devono guidare l'individuazione delle eccezioni e limitazioni. L'IA rende questi problemi ancora più urgenti: la generazione automatica di opere attraverso la combinazione di elementi preesistenti secondo pattern statistici pone il dubbio sulla loro qualificazione come opere originali o come meri derivati, mentre l'estrazione massiva di contenuti protetti per l'addestramento di modelli generativi apre la questione della liceità di tali pratiche rispetto alle norme sul diritto d'autore.

Di qui, la necessità di esaminare criticamente i tre principali confini che delimitano l'ambito del diritto d'autore: *i*) il requisito dell'originalità; *ii*) la distinzione idea-espressione; *iii*) la tensione tra monopolio privato e utilità sociale. Come si vedrà, nessuno di questi confini è tracciato in modo netto e definitivo, ma ciascuno rappresenta un equilibrio instabile che deve essere continuamente rinegoziato alla luce dei mutamenti tecnologici e sociali. L'IA generativa, con la sua capacità di produrre opere in quantità e a velocità senza precedenti, mette a dura prova questi equilibri tradizionali e richiede un ripensamento dei fondamenti stessi del sistema.

L'originalità è il criterio cardine che determina quali opere meritino di essere ricomprese entro la protezione del diritto d'autore e quali debbano rimanere nel dominio pubblico. Solo ciò che è originale – ossia ciò che presenta un carattere creativo sufficiente e non integra una mera riproduzione o una semplice imitazione di opere preesistenti – può essere oggetto di diritti esclusivi. Come è intuitivo, riuscire a individuare e a definire con precisione che cosa si intenda per originalità è tutt'altro che semplice. Sul punto, la giurisprudenza e la dottrina hanno via via elaborato molteplici concezioni, che, fra l'altro, non sono sempre compatibili tra loro<sup>24</sup>.

<sup>24</sup> Come è stato osservato: "l'originalità è un concetto sfuggente che assume significati diversi a seconda del contesto, del tipo di opera, e della tradizione giuridica di rife-

Una prima concezione – storicamente affermatasi soprattutto nel Regno Unito e in altri ordinamenti di *common law* prima dell’armonizzazione europea – identifica l’originalità con lo sforzo, l’abilità e il giudizio (*skill, labour and judgment*) impiegati dall’autore nella realizzazione dell’opera. Tale impostazione, tuttavia, non coincide con lo standard oggi affermato nel diritto statunitense, dove la giurisprudenza richiede la creazione indipendente e almeno un minimo grado di creatività, escludendo che il mero investimento di lavoro sia di per sé sufficiente. Secondo questo *standard*, anche opere prive di particolare merito estetico o intellettuale possono essere protette nel momento in cui hanno richiesto un investimento di tempo e di competenza<sup>25</sup>. Si tratta di una prima concezione che si avvicina molto alla ricostruzione lockiana del diritto d’autore e, in ragione della quale, quest’ultimo merita protezione perché ha sommato il proprio lavoro con il materiale preesistente. Tale *standard* è stato però criticato in quanto ritenuto eccessivamente permissivo e incline a confondere l’originalità con il mero investimento economico<sup>26</sup>.

Nell’Unione Europea, come si è fatto cenno, la giurisprudenza della Corte di Giustizia ha elaborato uno *standard* di originalità fondato sulla nozione di “creazione intellettuale propria dell’autore”. Ed in tal senso, soprattutto, nelle sentenze *Infopaq* e *Painer*, la Corte ha chiarito che un’opera è originale quando riflette le scelte libere e creative dell’autore, imprimendo così su di essa il suo “tocco” personale. Questo *standard* cerca di mediare tra la concezione anglosassone (che enfatizza lo sforzo) e quella continentale (che dà valenza all’espressione della personalità), ma solleva degli interrogativi circa il grado di libertà e di creatività richiesto, soprattutto in ordine a quante (e quali) scelte deve operare l’autore affinché l’opera possa essere considerata originale.

La questione diventa ancora più complessa nel momento in cui si considerano opere generate attraverso processi automatizzati o stocastici. In

rimento” (L. BENTLY, B. SHERMAN, *Intellectual Property Law*, Oxford, Oxford University Press, 2014).

<sup>25</sup> L’esempio classico è quello degli elenchi telefonici o delle raccolte di dati: pur non presentando alcuna creatività in senso proprio, essi sono originali nella misura in cui richiedono lavoro di compilazione, verifica e organizzazione.

<sup>26</sup> La Corte Suprema degli Stati Uniti, nella già citata sentenza *Feist* del 1991, ha rigettato espressamente la teoria dello *sweat of the brow*, affermando che “la Costituzione richiede un livello minimo di creatività, e sforzo, abilità e giudizio da soli non soddisfano tale requisito”. In modo particolare, secondo la Corte, un elenco telefonico compilato in ordine alfabetico non è originale poiché l’ordine alfabetico è una scelta ovvia e meccanica che non implica alcuna creatività. Questa posizione riflette una concezione dell’originalità come espressione di scelte creative individuali piuttosto che come risultato di mero lavoro.

questo caso, invero, la combinazione di elementi preesistenti secondo regole probabilistiche solleva il problema della qualificazione del risultato in termini di originalità. Nel dettaglio, se si guarda al piano del processo, mancano scelte creative nel senso ordinario, poiché l'algoritmo non decide consapevolmente come combinare gli elementi, ma applica meccanicamente funzioni matematiche. Diversamente, se ci si sofferma sul piano del risultato l'opera può presentare caratteristiche formalmente indistinguibili rispetto a quelle di opere create da dall'umano. Il nodo interpretativo – come è evidente – risiede nella determinazione dell'aspetto al quale attribuire la prevalenza nelle more della valutazione dell'originalità.

Alcuni autori hanno sostenuto che l'originalità dovrebbe essere valutata in base al risultato piuttosto che al processo: se l'opera apparisse nuova e non fosse copia di opere preesistenti, essa dovrebbe essere considerata originale indipendentemente da come è stata generata. Questa posizione, tuttavia, si scontra con il principio – consolidato nella giurisprudenza europea – che l'originalità presuppone scelte libere e creative, attributo che possiedono solo gli esseri umani. All'opposto, altri autori hanno proposto di riconoscere come originali le opere generate dall'IA solo quando vi sia un intervento umano significativo nella fase di concezione, di selezione o di *editing*, ma questa soluzione richiede di specificare quale grado di intervento sia sufficiente.

Un ulteriore problema riguarda la relazione tra originalità e novità. Un'opera può essere originale anche se simile a opere preesistenti, purché sia stata creata indipendentemente; viceversa, un'opera può essere nuova (ossia mai esistita prima) ma non essere affatto originale, può, infatti, accadere che ci si trovi di fronte a una combinazione banale o prevedibile di elementi già noti. La distinzione qui proposta – che è chiara quanto meno in astratto – diventa assai sfumata e labile nella pratica. La forte somiglianza tra due opere non consente di per sé di stabilire con certezza se essa derivi da copiatura o da creazione indipendente. Analogamente, la semplicità estrema di un'opera pone il dubbio sulla sua riconducibilità a genuine scelte creative, con conseguenti ricadute sulla sua proteggibilità.

L'IA complica ulteriormente questo quadro. I sistemi generativi addestrati su vasti *corpus* di opere protette producono *output* che, pur essendo statisticamente nuovi (nel senso che non replicano esattamente alcuna opera del training set), sono comunque determinati dalla distribuzione statistica iniziale del *corpus* di partenza. Sono originali opere che emergono da un processo di interpolazione e di ricombinazione di *pattern* estratti da milioni di creazioni umane? La risposta a questa domanda non è puramente tecnica, ma investe la concezione stessa di originalità e di creatività che si intende adottare nel diritto d'autore del futuro.

Una via percorribile – ad avviso di chi scrive – potrebbe essere quella di adottare un criterio normativo orientato alla funzione: l'originalità di un'opera generata con l'ausilio dell'IA andrebbe quindi valutata non in astratto, ma in relazione al grado di controllo creativo esercitato dall'operatore umano nella fase di concezione, di selezione e di revisione dell'*output*. Questo criterio funzionale-relazionale non dissolve il requisito dell'originalità, ma lo ricontestualizza: in pratica, esso riconosce che il tocco personale richiesto dalla giurisprudenza europea può manifestarsi anche attraverso la mediazione algoritmica, a condizione che vi sia una catena causale identificabile tra le scelte espressive dell'operatore e le caratteristiche dell'opera risultante.

Uno dei principi fondamentali del diritto d'autore è la distinzione tra idea ed espressione, secondo cui il diritto esclusivo copre solo le forme espressive attraverso cui le idee sono manifestate, ma non le idee stesse. Questo principio – enunciato per la prima volta nella giurisprudenza statunitense e poi recepito a livello internazionale – costituisce uno dei cardini che delimitano l'ambito della protezione e preservano la libertà di espressione e il progresso della conoscenza. Come affermato nella sentenza *Baker v. Selden*<sup>27</sup>: “la descrizione di un'arte in un libro non può essere l'oggetto di un *copyright*, per quanto sia originale tale descrizione; è solo l'espressione particolare della descrizione che appartiene all'autore”.

La *ratio* di questa distinzione è duplice. Da un lato, tutelare le idee significherebbe attribuire a chi per primo le formula un monopolio permanente sulla conoscenza, impedendo ad altri di svilupparle, criticarle o applicarle: questo contrasterebbe con il fine stesso del diritto d'autore, che è quello di promuovere il progresso culturale e scientifico. Dall'altro lato, le idee – a differenza delle espressioni – non sono il prodotto di un atto creativo individuale ma emergono spesso simultaneamente in contesti diversi, oppure sono il risultato di elaborazioni collettive progressive. In questi casi, attribuire a un singolo soggetto la proprietà di un'idea significherebbe misconoscere la natura sociale e collaborativa della produzione intellettuale.

Tuttavia, tracciare il confine tra idea ed espressione si rivela estremamente problematico. Nel caso di un romanzo, ad esempio, non è affatto chiaro se debbano qualificarsi come idea la sola trama generale o anche elementi più specifici quali la struttura narrativa, i personaggi e i dialoghi. Nella sentenza *Nichols v. Universal Pictures Corp.*<sup>28</sup>, il giudice Learned

<sup>27</sup> United States Supreme Court, *Baker v. Selden*, 101 U.S. 99 (1879).

<sup>28</sup> United States Court of Appeals, *Nichols v. Universal Pictures Corp.*, 45 F.2d 119 (1930).

Hand riconobbe che “è impossibile fissare il punto esatto in cui l’idea diventa espressione protetta” e propose una gradazione: più il livello di astrazione è elevato, più ci si avvicina all’idea non protetta; più si scende nei dettagli, più ci si avvicina all’espressione protetta. Ma questa formulazione, per quanto intuitiva, non fornisce un criterio operativo preciso<sup>29</sup>.

A tal proposito, è utile rammentare la teoria della documentalità elaborata da Maurizio Ferraris, in base alla quale, la formula “oggetto sociale = atto iscritto” non descrive la semplice registrazione di un contenuto preesistente: l’iscrizione è l’atto costitutivo attraverso cui un contenuto diventa oggetto sociale, entra nella realtà condivisa, acquista uno statuto accessibile. Se applicata alla distinzione dell’idea/espressione, tale prospettiva conduce a una conclusione di rilievo: ciò che il diritto d’autore protegge – vale a dire l’espressione – è precisamente ciò che, nella terminologia ferrarisiana, costituisce l’oggetto sociale; mentre l’idea non iscritta è, per definizione, un fatto bruto ontologicamente irrilevante per il diritto. Se così, la distinzione idea/espressione non rappresenta solo una scelta di politica normativa, ma corrisponde a una differenza ontologica profonda tra ciò che è iscritto – e quindi socialmente accessibile – e ciò che, invece, non lo è ancora.

Questa lettura ha conseguenze dirette sul problema dell’addestramento dei modelli generativi. Le opere su cui vengono addestrati i sistemi di IA sono oggetti sociali già iscritti: non mere idee fluttuanti, ma testi, immagini, partiture fissate su supporti. Interrogarsi se l’addestramento costituisca uso lecito o appropriazione illecita significa interrogarsi su quale statuto ontologico attribuire all’operazione di estrazione statistica compiuta su tali iscrizioni: essa rispetta il confine tra idea non protetta ed espressione protetta, o lo dissolve estraendo *pattern* che catturano al contemporaneo strutture semantiche e stilistiche delle opere iscritte? Come si dirà, la risposta a questa domanda non può essere puramente tecnica, ma richiede una presa di posizione sulla natura ontologica dei *pattern* statistici estratti.

<sup>29</sup> La difficoltà di tracciare il confine emerge con particolare evidenza nel *software* e nelle opere funzionali. Un algoritmo è un’idea oppure un’espressione? Il codice sorgente che implementa l’algoritmo è certamente espressione protetta, ma se due programmatori implementano lo stesso algoritmo in modi diversi, le due implementazioni, pur essendo espressioni distinte, potrebbero essere così simili da rendere difficile stabilire se vi sia stata copia dell’espressione o mera applicazione della stessa idea. La dottrina del *merger*, elaborata dalla giurisprudenza statunitense, stabilisce che quando vi è un solo modo o pochi modi per esprimere un’idea, l’espressione si fonde con l’idea e non può essere protetta; ma determinare quando questa situazione si verifichi richiede valutazioni discrezionali.

Un sistema generativo addestrato su un *corpus* di opere ne apprende tanto gli aspetti semantici – idee, temi, concetti – quanto quelli sintattici e stilistici, quali strutture linguistiche, registri espressivi e tecniche narrative. Ciò rende problematica la qualificazione giuridica del processo di addestramento: non è chiaro se esso costituisca un uso lecito delle opere, in quanto limitato all'estrazione di idee non proteggibili, ovvero un'appropriazione indebita anche delle forme espressive tutelate dal diritto d'autore. La questione è particolarmente controversa. Alcuni sostengono che l'addestramento dei modelli generativi costituisca un uso trasformativo e non consumativo delle opere, analogo alla lettura umana. Altri obiettano che l'IA, a differenza dell'essere umano, memorizza e ricombina *pattern* in modo meccanico, e che in molti casi gli output generati replicano stili o frasi tratte dalle opere di addestramento, configurando così una violazione del diritto esclusivo.

La giurisprudenza su questo punto è ancora agli inizi. Alcune cause intentate da autori e editori contro sviluppatori di sistemi di IA generativa sono pendenti negli Stati Uniti e in Europa, e le loro decisioni potrebbero ridefinire l'interpretazione della distinzione idea-espressione nell'era algoritmica.

L'indagine qui condotta si propone di contribuire a questo dibattito a partire da un criterio di fondo, vale a dire quello secondo il quale: la liceità dell'addestramento su corpo protetto dovrebbe essere valutata in funzione della trasformazione prodotta dall'operazione di allenamento rispetto alle opere di partenza. Usi trasformativi – quelli che estraggono competenze generalizzabili senza replicare espressioni specifiche – andrebbero distinti dagli usi appropriativi, che invece riproducono o ricostruiscono elementi stilistici individuali delle opere di addestramento. Tale distinzione, che riecheggia la dottrina del *fair use trasformativo* elaborata dalla giurisprudenza statunitense, necessita tuttavia di essere adattata al contesto dell'UE attraverso le eccezioni di *text and data mining* previste dalla Direttiva Copyright 2019/790.

Come è intuitivo, il diritto d'autore è attraversato da una tensione costitutiva tra due esigenze apparentemente contrapposte: per un verso, l'attribuzione all'autore di diritti esclusivi che incentivino la creazione; per l'altro, la garanzia dell'accesso pubblico alla conoscenza e alla cultura che permetta il progresso collettivo. Questa tensione non è un difetto del sistema, ma ne costituisce la struttura portante: il diritto d'autore realizza un compromesso tra interesse privato e utilità sociale, riconoscendo che entrambi sono necessari per il funzionamento di una società democratica e culturalmente vivace.

La metafora del bilanciamento tra incentivo e accesso, però, rischia di essere fuorviante se interpretata in senso statico. Non si tratta semplicemente di trovare un punto di equilibrio ottimale tra due valori dati, ma di riconoscere che i due poli sono dinamicamente interconnessi: troppa protezione soffoca la creatività futura limitando l'accesso al patrimonio culturale su cui i nuovi autori devono costruire; troppo poca protezione disincentiva la creazione riducendo la disponibilità di opere<sup>30</sup>.

Storicamente, questo bilanciamento è stato realizzato attraverso diversi meccanismi. Innanzitutto, la limitazione temporale della protezione: l'opera è protetta per un periodo determinato (attualmente vita dell'autore più settant'anni nell'UE e negli USA), dopodiché entra nel dominio pubblico e può essere liberamente utilizzata. In secondo luogo, le eccezioni e limitazioni: anche durante il periodo di protezione, determinati usi (citazione, critica, uso privato, finalità didattiche, parodia) sono permessi senza autorizzazione. In terzo luogo, la distinzione idea-espressione: come si è visto, solo le forme espressive sono protette, mentre le idee rimangono liberamente utilizzabili.

Tuttavia, l'equilibrio realizzato da questi meccanismi è stato progressivamente messo in discussione. La durata della protezione è stata ripetutamente estesa nel corso del XX secolo, in particolare negli Stati Uniti con il *Sonny Bono Copyright Term Extension Act* del 1998, che ha prolungato il *copyright* di ulteriori vent'anni. Questa estensione è stata criticata da molti autori e attivisti come un'appropriazione indebita del dominio pubblico a vantaggio delle grandi *corporation* dell'*entertainment*. Come ha argomentato Lawrence Lessig nel suo libro *Free Culture*, "l'estensione del termine di protezione non incentiva la creazione di opere già esistenti, ma serve solo a proteggere investimenti passati a scapito dell'innovazione futura"<sup>31</sup>.

Analogamente, le eccezioni e limitazioni sono state progressivamente ridotte o rese più difficilmente applicabili attraverso misure tecnologiche di protezione (*DRM*) e disposizioni contrattuali. Il *digital lock* e le licenze d'uso, pur essendo strumenti legittimi di protezione dei diritti, possono essere impiegati per impedire usi che sarebbero leciti secondo le eccezioni previste dalla legge. Questa privatizzazione del diritto d'autore, operata attraverso il contratto e la tecnologia, rischia di svuotare di significato il bilanciamento pubblico stabilito dal legislatore.

<sup>30</sup> Come ha osservato LITMAN: "il *copyright* deve essere concepito non come un sistema di protezione proprietaria ma come un'istituzione che facilita la produzione e la circolazione culturale" (J. LITMAN, *Digital Copyright*, Amherst, Prometheus Books, 2001).

<sup>31</sup> L. LESSIG, *Free Culture: How Big Media Uses Technology and the Law to Lock Down Culture and Control Creativity*, New York, Penguin Press, 2004.

Dal punto di vista filosofico, la tensione tra monopolio privato e utilità sociale riflette un più ampio dibattito sulla natura della proprietà intellettuale e sulla legittimità delle forme di appropriazione della conoscenza. Autori ispirati alla tradizione libertaria, come Robert Nozick, sostengono che i diritti di proprietà intellettuale sono moralmente equivalenti a quelli sulla proprietà materiale e devono essere protetti con pari rigore. Altri, influenzati dal pensiero socialista o comunitarista, argomentano che la conoscenza è un bene comune che non dovrebbe essere appropriabile privatamente, o che l'appropriazione deve essere strettamente limitata e condizionata al servizio del bene pubblico.

L'intelligenza artificiale introduce nuovi elementi in questa tensione. Da un lato, i sistemi generativi abbassano drasticamente i costi di produzione di contenuti culturali, rendendo meno necessario il meccanismo dell'incentivo economico: se chiunque può generare testi, immagini o musiche con pochi clic, ha ancora senso attribuire monopoli temporanei per incentivare la creazione? Dall'altro lato, lo sviluppo di sistemi di IA richiede investimenti enormi in ricerca, dati e infrastrutture computazionali, e le imprese tecnologiche sostengono che senza protezione giuridica adeguata tali investimenti non sarebbero sostenibili.

Inoltre, l'IA pone il problema del data mining e del *text and data mining* (TDM): l'estrazione automatizzata di informazioni da grandi *corpus* di opere protette per finalità di ricerca, addestramento di modelli o altre applicazioni. Molte giurisdizioni hanno introdotto eccezioni specifiche per il TDM, riconoscendo che impedire tali usi ostacolerebbe l'innovazione scientifica e tecnologica. Tuttavia, i titolari dei diritti sostengono che il TDM su scala industriale costituisce un uso commerciale che dovrebbe richiedere autorizzazione e compenso.

L'indagine qui prospettata affronta queste questioni cercando di individuare un nuovo equilibrio che tenga conto delle specificità dell'IA generativa. L'obiettivo non è preservare acriticamente il bilanciamento tradizionale, che potrebbe rivelarsi inadeguato nel nuovo contesto tecnologico, ma nemmeno abbandonare i principi fondamentali che hanno giustificato storicamente il diritto d'autore. Si tratta, piuttosto, di elaborare soluzioni flessibili e graduali che permettano di distinguere tra usi legittimi dell'IA (quelli che promuovono l'innovazione senza danneggiare ingiustamente gli autori) e usi illeciti (quelli che si appropriano del lavoro creativo altrui senza offrire contropartite adeguate).

In questa prospettiva, possono individuarsi tre criteri normativi di orientamento per il legislatore e l'interprete: *i)* il criterio della proporzionalità della tutela, in base al quale la protezione accordata alle opere

generate o co-generate con l'IA deve essere proporzionale al grado di intervento creativo umano effettivamente esercitato; *ii*) il criterio della non regressione del dominio pubblico, che impone di valutare le nuove forme di protezione in termini di impatto sull'accesso collettivo alla conoscenza; *iii*) il criterio della responsabilità distribuita, seguendo il quale, in assenza di un autore umano identificabile, i diritti sulle opere di IA devono essere attribuiti in modo da non creare vuoti di responsabilità per usi illeciti o dannosi. Questi criteri non pretendono di risolvere le tensioni strutturali del sistema, eppure offrono parametri di giudizio coerenti con i valori fondamentali che il diritto d'autore è storicamente chiamato a tutelare.

PARTE SECONDA

*IA e crisi del paradigma giuridico*



## CAPITOLO III

### LA SFIDA ONTOLOGICA DELL'IA GENERATIVA AL DIRITTO D'AUTORE

SOMMARIO: 1. L'IA generativa come nuovo attore del processo creativo. – 2. La crisi del concetto di originalità nell'era algoritmica. – 3. Il problema dell'intenzionalità nelle creazioni generate da IA. – 4. La dimensione temporale della creatività artificiale.

#### 1. *L'IA generativa come nuovo attore del processo creativo*

I due capitoli precedenti hanno ricostruito i fondamenti filosofici e giuridici del diritto d'autore moderno, evidenziando come esso si radichi sulla visione della creazione come attività eminentemente ed esclusivamente umana, in quanto manifestazione di facoltà cognitive ed emotive che solo le persone possiedono. Come è evidente, l'IA generativa – intesa come insieme di sistemi computazionali capaci di produrre autonomamente testi, immagini, musiche e altri contenuti – rappresenta una sfida radicale a tale assunto. Per la prima volta nella storia, infatti, ci troviamo di fronte a una situazione in cui entità non umane dimostrano capacità creative che – almeno sul piano fenomenico – sembrano indistinguibili da quelle umane. Ovviamente, tale fenomeno richiede non solo un adattamento tecnico delle norme esistenti, ma anche – e soprattutto – un profondo ripensamento filosofico della natura della creatività e del ruolo che l'autore svolge all'interno della società e del sistema giuridico.

In questo capitolo si guarderà alla sfida ontologica che l'IA generativa pone al diritto d'autore, concentrandosi in via prevalente su quattro dimensioni fondamentali: *i)* l'emergere dell'IA come attore del processo creativo; *ii)* la crisi del concetto di originalità quando applicato a produzioni algoritmiche; *iii)* il problema dell'intenzionalità nelle creazioni generate da sistemi privi di coscienza; *iv)* la dimensione temporale della creatività artificiale, caratterizzata da velocità e serialità senza precedenti. Ciascuna di queste dimensioni mette in discussione un pilastro dell'impianto tradizionale del diritto d'autore, rendendo necessaria quella rifondazione teorico-pratica che l'indagine mira a incentivare.

Prima ancora di affrontare le questioni filosofiche e giuridiche sollevate dall'IA generativa, è indispensabile comprendere che cosa sia tecni-

camente tale tecnologia e come funzioni. L'espressione "intelligenza artificiale generativa" designa una classe di sistemi computazionali basati su algoritmi di *machine learning* che, addestrati su vasti *corpus* di dati, sono capaci di generare nuovi contenuti – testi, immagini, audio, video – che presentano caratteristiche statisticamente coerenti con i dati di addestramento ma non sono copie esatte di alcun elemento del *corpus*. Tali sistemi si distinguono dalle precedenti generazioni di IA per la loro capacità di produrre *output* originali piuttosto che limitarsi a riconoscere *pattern* o classificare dati esistenti.

Il funzionamento dell'IA generativa si basa su architetture neurali profonde, in particolare su modelli *transformer* per il testo (come GPT, sviluppato da OpenAI) e su diverse architetture generative per le immagini, tra cui i modelli di diffusione e altri sistemi proprietari di generazione visuale (come DALL·E, Midjourney e Stable Diffusion). Senza entrare nei dettagli tecnici – che esulano dallo scopo di questa trattazione – è importante comprenderne i principi generali.

Un modello generativo viene addestrato su un grande *dataset* – ad esempio, milioni di coppie immagine-descrizione testuale – e impara a riconoscere le correlazioni statistiche tra i diversi elementi. Durante l'addestramento, il modello ottimizza i propri parametri interni (che possono essere miliardi) in modo da minimizzare l'errore di predizione. Una volta addestrato, il modello può generare nuovi contenuti a partire da un *input* fornito dall'utilizzatore (il cosiddetto *prompt*). Ad esempio, nel momento in cui l'utente fornisce la descrizione testuale "un gatto che suona il pianoforte in stile impressionista", il sistema è in condizione di generare un'immagine che corrisponde a tale descrizione, combinando elementi appresi durante l'addestramento (forme di gatti, pianoforti, caratteristiche dello stile impressionista) in una configurazione nuova. Fondamentalmente, il sistema non sta semplicemente recuperando e combinando immagini del *dataset* di addestramento, ma sta interpolando nello spazio latente delle rappresentazioni apprese, generando così *output* che non esistevano nel *corpus* di addestramento originario.

Questa capacità di interpolazione e di generazione di novità impone un'immediata riflessione filosofica e giuridica sulla natura dell'*output*. Occorre, innanzitutto, chiarire in quale misura esso possa definirsi realmente "nuovo", data la sua derivazione statistica dai dati di addestramento. Parallelamente, è necessario stabilire se l'apprendimento di *pattern* da opere protette configuri o meno una violazione dei diritti degli autori originali.

Sul piano della titolarità, si pone il problema di individuare il soggetto responsabile dell'*output* generato, ripartendo gli oneri tra il program-

matore del sistema, l'utente che ha fornito il *prompt* e l'azienda proprietaria del modello, oppure di considerare la possibilità di ritenere assente la responsabilità individuale definita. La risoluzione di tali nodi problematici richiede una distinzione analitica tra le diverse fasi del processo e i molteplici attori coinvolti nella creazione mediata da IA.

Una prima fase è quella dello sviluppo del modello: ricercatori e ingegneri progettano l'architettura neurale, selezionano i dati di addestramento, definiscono gli obiettivi di ottimizzazione e addestrano il modello. Questo momento richiede competenze tecniche elevate, investimenti computazionali significativi e scelte progettuali che influenzano le capacità e determinano i limiti del sistema risultante.

Una seconda fase è quella dell'utilizzo: un utente interagisce con il sistema fornendo *input* (mediante la formulazione di *prompt*) e ottenendo *output*. Questo frangente può essere più o meno creativo: un utente esperto può interagire molteplici volte, raffinando i *prompt* e selezionando tra diversi *output* quello più soddisfacente. Diversamente, un utente occasionale può limitarsi a inserire una descrizione generica e successivamente ad accettare il primo risultato proposto.

Una terza fase, spesso trascurata ma giuridicamente rilevante, è quella del *post-processing*: l'*output* generato automaticamente può essere successivamente modificato, integrato con elementi creati manualmente o incorporato in opere composite. Ad esempio, un illustratore può generare uno sfondo con *Midjourney* e poi aggiungere manualmente personaggi disegnati a mano; uno scrittore può usare *GPT* per generare una prima bozza e poi riscriverla estensivamente. In questi casi, l'opera finale è il risultato di un processo ibrido in cui l'IA contribuisce solo parzialmente e il contributo umano può essere più o meno sostanziale.

Queste distinzioni sono cruciali per valutare le diverse soluzioni giuridiche possibili. Se si ritiene che la creatività risieda nella fase di sviluppo del modello, allora sono i programmatori a dover essere considerati gli autori di tutte le opere generate dal loro sistema: soluzione evidentemente assurda data la quantità e la varietà degli *output*. Se, invece, si ritiene che la creatività risieda nella fase di utilizzo, allora sono gli utenti a dover essere riconosciuti come gli autori, ma – in questo caso – occorre stabilire quale grado di intervento sia sufficiente. Se, ancora, si ritiene che la creatività sia distribuita tra le diverse fasi, allora è necessario elaborare categorie di co-autorialità o di opere composite. Come si vedrà nei paragrafi successivi, nessuna di queste soluzioni è priva di problemi.

Un ulteriore aspetto tecnico rilevante riguarda la questione del *data mining* e dell'addestramento. I modelli generativi richiedono *dataset*

enormi per funzionare efficacemente, nel caso di GPT-4 la documentazione pubblica di OpenAI indica che il modello è stato addestrato su un *web-scale corpus* composto da dati pubblicamente disponibili e da dati concessi in licenza, senza però rendere nota in modo analitico la dimensione complessiva del *training set*. Molte di queste opere sono protette da *copyright*. Ne deriva, pressoché immediatamente, un interrogativo cruciale: l'addestramento del modello su tali opere costituisce una violazione dei diritti esclusivi degli autori? Gli sviluppatori di IA sostengono che si tratta di un uso trasformativo e non consumativo, analogo a quello che si ha tramite la lettura umana o la ricerca accademica. I titolari dei diritti, di contro, obiettano che si tratta di un'appropriazione massiva di valore economico senza un'autorizzazione o un compenso. La giurisprudenza su questo punto è ancora in formazione. Negli Stati Uniti sono tuttora pendenti controversie significative – tra cui *Andersen v. Stability AI*<sup>1</sup> – che contestano la legittimità dell'addestramento dei modelli su opere protette in assenza di licenza. In Europa, la Direttiva (UE) 2019/790 ha introdotto, da un lato, un'eccezione obbligatoria per il *text and data mining* a fini di ricerca scientifica (art. 3) e, dall'altro lato, un'eccezione o limitazione più generale per il *text and data mining* su contenuti accessibili lecitamente (art. 4), fatta salva la possibilità per i titolari dei diritti di riservarne espressamente l'uso, ad esempio con mezzi *machine-readable* nel caso di contenuti resi disponibili online. Proprio l'applicabilità di questo quadro all'addestramento di modelli commerciali rimane, allo stato, controversa.

L'indagine, qui condotta, punta ad affrontare questa questione prospettando un bilanciamento fra l'interesse degli autori alla tutela delle proprie e quello sociale (generale) all'innovazione tecnologica e alla ricerca.

Comprendere gli aspetti tecnici del funzionamento dell'IA generativa è condizione necessaria ma non sufficiente a determinarne lo statuto ontologico e giuridico. È necessario riflettere filosoficamente circa la possibilità che tali sistemi possano essere considerati agenti – ossia soggetti dotati di capacità di azione autonoma – o se, invece, rimangano meri strumenti attraverso i quali gli esseri umani agiscono. La questione dell'*agency* è centrale perché il diritto attribuisce responsabilità e riconosce diritti solo in capo agli agenti: in pratica solamente chi agisce può essere ritenuto responsabile delle conseguenze delle proprie azioni e può vantare diritti sulle opere che produce. Ragion per cui, se l'IA non è un agente, allora essa non può essere

<sup>1</sup> *Andersen et al. v. Stability AI Ltd. et al.*, United States District Court for the Northern District of California, No. 3:23-cv-00201-WHO (W.H. Orrick, J.), azione depositata il 13 gennaio 2023, procedimento in corso al 1° gennaio 2026.

né autrice, né responsabile; al contrario, se essa è equiparabile a un agente, occorre ripensare le categorie giuridiche tradizionali.

Nella filosofia dell'azione, un agente è tipicamente caratterizzato da tre proprietà fondamentali: *i*) l'intenzionalità (agisce per ragioni e scopi); *ii*) l'autonomia (le sue azioni non sono completamente determinate da cause esterne); *iii*) la razionalità (sceglie mezzi appropriati per i suoi fini). Gli individui (maggiorenni e dotati di capacità di intendere e di volere) sono paradigmaticamente ritenuti degli agenti in questo senso: essi decidono consapevolmente di compiere determinate azioni, valutano le possibili alternative, stabiliscono e perseguono obiettivi. Gli animali (non umani) presentano gradi variabili di *agency*: i mammiferi superiori mostrano comportamenti intenzionali e capacità di scelta, mentre organismi più semplici agiscono principalmente per riflessi innati. Le macchine tradizionali – un orologio, un'automobile, un computer che esegue un programma – non sono agenti perché non sono dotati di finalità e di scopi propri né, tanto meno, scelgono autonomamente.

L'IA generativa si colloca in una posizione a suo modo intermedia e per così dire ambigua rispetto a queste categorie. Da un lato, essa non possiede intenzionalità nel senso fenomenologico: un sistema di IA non 'desidera' generare un'immagine, non 'prova' soddisfazione o compiacimento nel riuscire a produrre un'opera, non 'si preoccupa' della qualità dell'*output*. La sua attività è sostanzialmente il risultato di operazioni matematiche determinate dall'architettura del modello e dai parametri appresi durante l'addestramento. Da un altro lato, però, l'IA denota comportamenti che, visti dall'esterno, appaiono intenzionali e autonomi: quando *GPT* genera un testo coerente e pertinente alla richiesta dell'utente, sembra comprendere la domanda e scegliere la risposta più appropriata.

Sul versante scettico, tanto l'argomento di Searle (riferito alla *Stanza cinese*)<sup>2</sup> quanto la critica di Dreyfus all'IA (come sistema privo di radicamento esistenziale)<sup>3</sup> convergono nell'affermare che i sistemi computazio-

<sup>2</sup> J. SEARLE, *Menti, cervelli e programmi. Un dibattito sull'intelligenza artificiale*, trad. it., Milano, Clup, 1984; ID., *The Rediscovery of the Mind*, Cambridge (MA), MIT Press, 1992.

<sup>3</sup> La critica di HUBERT DREYFUS all'intelligenza artificiale come sistema privo di radicamento esistenziale e corporeo è sviluppata in H.L. DREYFUS, *What Computers Can't Do: A Critique of Artificial Reason*, New York, Harper & Row, 1972, e ripresa e aggiornata in ID., *What Computers Still Can't Do: A Critique of Artificial Reason*, Cambridge (MA), MIT Press, 1992. La tesi centrale è che l'intelligenza umana non sia riducibile a elaborazione formale di regole simboliche, poiché dipende in modo costitutivo dall'incarnazione in un corpo che si confronta praticamente con il mondo ("*embodied coping*"). In tale prospettiva, sistemi computazionali privi di corpo e di esperienza vissuta non possono raggiungere una comprensione autentica, qualunque sia la loro sofisticazione algoritmica.

nali, manipolando simboli senza comprenderli e agendo senza un corpo che si confronti praticamente con il mondo, sono strutturalmente esclusi dalla possibilità di vedersi riconosciuta un'agency autentica<sup>4</sup>. Applicata al diritto d'autore, questa stessa posizione implica una conseguenza precisa: un sistema che non "sa" che cosa sta producendo e non vive interiormente il processo creativo non può essere considerato autore, perché è privo di quella consapevolezza espressiva che giustifica la protezione giuridica. L'output algoritmico sarebbe, in questa prospettiva, analogo al risultato di una forza naturale: formalmente e fattualmente rilevante, ma concettualmente privo dello statuto di opera in senso giuridico.

Sul versante opposto, la posizione intenzionale di Dennett<sup>5</sup> propone di adottare un criterio funzionale e pragmatico: un sistema è un agente – e può quindi essere trattato come tale – quando è utile e predittivamente efficace interpretarlo come se avesse credenze e desideri, indipendentemente dalla sua consapevolezza/autocomprensione interna. Questa prospettiva non ambisce a chiamare in causa o a risolvere la questione metafisica della coscienza: nel caso di specie ciò che si ritiene sufficiente è che il sistema si comporti "come" un agente. I possibili risvolti in tema di diritto d'autore sono evidenti: se l'agency è una questione di grado (in termini che potremmo definire quantitativi), allora anche la tutela autorale può essere calibrata sul grado di agency esibita. Un sistema che si comporta come agente creativo in misura ridotta – perché il suo apporto alla creazione è limitato o interamente predeterminato dal training – meriterebbe una forma di protezione ridotta o condizionata, e, in ogni caso, comunque diversa rispetto alla tutela riservata all'autore umano.

La posizione intermedia assunta da Floridi in ordine all'agente artificiale in senso debole prospetta lo strumento concettuale per tradurre in termini concreti e operativi questa stessa gradazione. Se l'IA è un agente debole – capace di percepire e agire sull'ambiente, ma privo di coscienza morale – allora le opere che genera sono il prodotto di un'agency parzia-

Per un'applicazione di queste tesi al dibattito contemporaneo sull'IA si veda anche H.L. DREYFUS, S.E. DREYFUS, *Mind over Machine: The Power of Human Intuition and Expertise in the Era of the Computer*, New York, Free Press, 1986.

<sup>4</sup> Sul nesso tra opacità dei processi algoritmici e impossibilità di ricostruire dall'esterno intenzioni o motivazioni genuine, si veda G. GOMETZ, "Intelligenza artificiale, profilazione e nuove forme di discriminazione", in *Teoria e Storia del Diritto Privato*, n. spec., 2022, pp. 1-38, ove si osserva che le IA basate su *machine learning* non sono in grado di presentare alcun resoconto delle ragioni che le hanno determinate a produrre un certo output, operando secondo un approccio "meramente statistic", che esclude qualsiasi forma di comprensione concettuale assimilabile a quella umana.

<sup>5</sup> D.C. DENNETT, *The Intentional Stance*, Cambridge, (MA), MIT Press, 1987.

le. Il diritto d'autore tradizionale è stato pensato per tutelare il risultato di un'*agency* piena: scelte libere, consapevoli, espressive di una personalità. Riservare la stessa tutela a produzioni di *agency* parziale significa applicare una categoria progettata per una determinata situazione a una situazione fondamentalmente diversa. Con questo lavoro ci si propone di affrontare questa asimmetria non mediante l'esclusione totale delle opere algoritmiche da uno scenario di protezione giuridica, ma ricorrendo piuttosto ad un sistema di tutele differenziate, che siano calibrate sull'effettivo grado di controllo creativo umano di volta in volta esercitato nel processo generativo.

Dal punto di vista giuridico, la questione dell'*agency* dell'IA ha conseguenze dirette. Se l'IA non è un agente, allora le sue azioni devono essere imputate a qualcun altro, che può essere: il programmatore, l'utilizzatore o il proprietario del sistema. Se, diversamente, l'IA è un agente, allora essa potrebbe essere considerata l'autrice delle opere che genera e la responsabile delle violazioni eventualmente commesse.

L'attribuzione all'IA della qualifica di agente solleva problemi pratici di complessa risoluzione. Sul piano dei diritti morali, occorre definire le modalità del loro eventuale esercizio da parte di un'entità artificiale, mentre sul fronte risarcitorio è necessario individuare i soggetti chiamati a rispondere patrimonialmente per i danni causati dalle sue azioni. Più radicalmente, si pone il problema di giustificare l'attribuzione di diritti e responsabilità a entità che risultano strutturalmente prive di interessi propri da tutelare.

L'indagine condotta affronta questa tensione proponendo un approccio pragmatico e graduato. Anziché decidere in astratto se l'IA sia o non sia un agente, si tratta di valutare caso per caso quale grado di autonomia e di capacità decisionale presenti un sistema specifico e di attribuire conseguenze giuridiche proporzionate a tale grado. Un sistema che genera *output* completamente deterministici a partire da *input* umani dettagliati presenta un'*agency* minima e dovrebbe essere trattato come un mero strumento. All'opposto, un sistema capace di apprendimento continuo e di operare scelte non previste dai programmatori presenta un'*agency* maggiore e potrebbe ricevere una qualificazione giuridica diversa. Questa flessibilità concettuale permette di evitare dogmatismi filosofici pur mantenendo la praticabilità giuridica.

La questione filosofica dell'*agency* dell'IA si traduce in una domanda giuridica concreta, ovverosia: quale qualificazione deve ricevere l'IA generativa nel contesto del diritto d'autore? Le opzioni principali sono tre: *i*) considerarla alla stregua di uno strumento nelle mani dell'essere umano; *ii*)

riconoscerla come autrice autonoma; *iii*) elaborare una categoria *sui generis* che riconosca la sua specificità senza forzarla e senza ricondurla a tutti i costi entro le categorie tradizionali. Ciascuna di queste opzioni presenta vantaggi e svantaggi. E, ovviamente, la scelta tra tali alternative dipende da quale equilibrio si intenda realizzare tra i diversi valori in gioco.

La prima opzione – vale a dire quella di considerare l’IA come uno strumento – è quella che meglio preserva la continuità con il sistema tradizionale del diritto d’autore. Secondo questa prospettiva, infatti, l’IA generativa è analoga alle tecnologie precedenti come la fotocamera, il sintetizzatore musicale o il software di grafica. In pratica, non sarebbe altro che uno strumento in grado di amplificare le capacità creative umane ma inabile a creare del tutto autonomamente nel senso proprio del termine. L’autore dell’opera generata con l’ausilio dell’IA sarebbe, dunque, l’essere umano che ha utilizzato il sistema, colui che ha fornito l’*input*, chi ha selezionato l’*output* ed eventualmente che poi ha anche modificato il risultato. Questa soluzione ha il merito della semplicità e della familiarità. In modo particolare, tale opzione non richiede modifiche legislative significative e può essere applicata attraverso l’interpretazione evolutiva delle norme esistenti.

Tuttavia, la qualificazione dell’IA come mero strumento incontra considerevoli difficoltà nel momento in cui il contributo umano sia minimo. Se un utente inserisce un *prompt* generico (genera un paesaggio) e accetta il primo *output* senza operare particolari modifiche, in che senso può dirsi davvero l’autore dell’opera risultante? La sua attività non sembra qualitativamente diversa da quella di chi aziona una *slot machine*: fornisce un *input* casuale e ottiene un risultato determinato da processi su cui non ha nessun tipo di controllo<sup>6</sup>. Inoltre, la qualificazione strumentale diventa vieppiù problematica ora che l’IA è capace di apprendimento continuo e di evoluzione autonoma. I sistemi di *reinforcement learning*, ad esempio, migliorano le proprie prestazioni attraverso l’interazione con l’ambiente senza supervisione umana diretta. In questi casi, gli *output* generati riflettono, non solo le scelte dei programmatori e degli utenti, ma anche i *pattern* appresi autonomamente dal sistema. Attribuire la paternità degli esiti agli esseri umani coinvolti significa ignorare quello che è il contributo effettivo dell’IA al processo creativo.

<sup>6</sup> Come ha osservato Samuelson, “se consideriamo autori coloro che semplicemente premono un pulsante e ottengono un’opera, stiamo svilendo il concetto stesso di autorialità creativa” (P. SAMUELSON, “Generative AI Meets Copyright”, in *Science*, 381(6654), 2023, pp. 1293-1294).

La seconda opzione – cioè quella di riconoscere l'IA come autrice – è stata proposta da alcuni teorici come soluzione coerente rispetto al problema delle opere generate autonomamente. In particolare, se l'IA produce opere originali, allora dovrebbe essere riconosciuta come autrice e titolare dei diritti, analogamente a quanto avviene per le persone giuridiche nel diritto societario. Ryan Abbott, ad esempio, ha difeso questa posizione argomentando che “negare la protezione alle opere generate da IA semplicemente perché non vi è un autore umano significherebbe sacrificare l'incentivo all'innovazione tecnologica senza giustificazione sufficiente”<sup>7</sup>.

Tuttavia, anche attribuire personalità giuridica e autorialità all'IA solleva considerevoli problemi teorici e pratici. Vien da sé: l'IA non ha interessi propri da tutelare, non può esercitare autonomamente i diritti morali e non può essere ritenuta responsabile per violazioni del *copyright*. Riconoscerla come autrice significherebbe dunque creare una vera e propria finzione giuridica che altri (proprietari del sistema, utenti) utilizzerebbero per i propri scopi, con possibili effetti distorsivi sul mercato e sulla produzione culturale. Inoltre, questa soluzione contrasterebbe con il principio – consolidato in molte giurisdizioni – che solo gli esseri umani possono essere ritenuti autori originari.

La terza opzione – ossia quella di elaborare una categoria *sui generis* – cerca di evitare i problemi delle vie precedenti riconoscendo che l'IA generativa non è né un mero strumento né un autore in senso pieno, ma costituisce un fenomeno nuovo che richiede categorie inedite. Questa opzione si ispira ai precedenti storici in cui il legislatore ha creato forme di tutela specifiche per fenomeni che non rientravano nelle categorie tradizionali, fra questi: i diritti *sui generis* per le banche dati (Direttiva 96/9/CE), i diritti connessi per le prestazioni degli artisti interpreti, i diritti sulle fotografie non creative in alcuni ordinamenti.

Una forma di tutela *sui generis* per le opere generate dall'IA potrebbe prevedere una protezione ridotta rispetto al *copyright* pieno, come ad esempio: una durata più breve, dei diritti patrimoniali ma non morali, delle limitazioni più ampie. La titolarità potrebbe essere attribuita all'impresa che ha sviluppato e commercializzato il sistema, riconoscendo così l'investimento tecnologico senza equiparare l'IA all'autore umano. Questa soluzione permetterebbe di incentivare l'innovazione senza svilire il concetto di autorialità creativa e evitando di sacrificare la specificità della creatività umana.

<sup>7</sup> R. ABBOTT, “I Think, Therefore I Invent: Creative Computers and the Future of Patent Law”, in *Boston College Law Review*, 57(4), 2016, pp. 1079-1126.

Tuttavia, anche la soluzione *sui generis* presenta difficoltà. Innanzitutto, occorre tracciare una distinzione chiara tra le opere create con l'ausilio dell'IA (che riceverebbero protezione piena) e le opere generate autonomamente dall'IA (che riceverebbero protezione *sui generis*). Si tratta di una distinzione, che, come si è visto, è assai sfumata: vi è un *continuum* di situazioni intermedie in cui il contributo umano e quello algoritmico si intrecciano e si fondono in modi complessi. In secondo luogo, una tutela differenziata rischia di creare incertezza giuridica e contenziosi circa la qualificazione delle opere, con costi transattivi elevati.

Alla luce di ciò, si propone un approccio flessibile e pluralista che non privilegia una soluzione univoca ma che cerca di individuare criteri per applicare trattamenti differenziati a situazioni diverse. L'idea guida è che non tutte le opere che coinvolgono l'IA devono ricevere lo stesso trattamento giuridico, ma che, al contrario, la tutela deve essere calibrata in base a diversi fattori quali: il grado di intervento umano creativo; l'autonomia effettiva del sistema; la trasparenza del processo generativo; l'impatto sugli autori umani e sul mercato culturale. Questo approccio richiede un'analisi caso per caso che sia guidata da principi normativi chiari, nell'intento di evitare tanto di sfociare in un formalismo rigido quanto di incappare in un'arbitrarietà discrezionale.

## 2. *La crisi del concetto di originalità nell'era algoritmica*

Il requisito dell'originalità, come si è accennato, costituisce il criterio cardine che determina quali opere meritino la protezione del diritto d'autore. Solo ciò che presenta un carattere creativo sufficiente, e che è espressione di scelte libere e creative da parte dell'autore, può essere oggetto di diritti esclusivi. Tuttavia, l'applicazione di questo criterio alle opere generate dall'IA solleva notevoli criticità sul profilo teorico e pratico. In particolare, se l'originalità presuppone la presenza di scelte consapevoli da parte di un soggetto creatore, come può essere ritenuta realmente originale un'opera prodotta da processi algoritmici deterministici o stocastici? E, ancora, se l'originalità viene valutata in base al risultato piuttosto che al processo, non si rischia di proteggere opere prive di quel carattere personale che ha storicamente giustificato la tutela autorale?

In questo senso, appare più che mai necessario, soffermarsi a riflettere sulla crisi del concetto di originalità nell'era algoritmica da tre prospettive diverse prospettive, fra loro, complementari. Innanzitutto, è opportuno procedere a una decostruzione filosofica del requisito dell'originalità, mostrando come tale nozione incorpori assunzioni circa la natura della creati-

vità che l'IA mette radicalmente in discussione. In secondo luogo, bisogna valutare se l'originalità possa o meno essere il prodotto di processi stocastici – ossia di combinazioni probabilistiche di elementi preesistenti – senza che vi sia intervento di scelte consapevoli. Infine, è necessario guardare alla dimensione della serialità creativa nell'epoca della riproducibilità algoritmica, interrogandosi su come il diritto d'autore debba/possa regolare fenomeni di produzione massiva di opere formalmente originali.

In realtà, come si vedrà, nessuna delle concezioni “tradizionali” di originalità risulta pienamente adeguata a dar conto delle varie espressioni della creatività algoritmica. L'idea, quindi, è quella di puntare ad elaborare criteri nuovi che permettano di distinguere tra opere meritevoli di tutela e mere produzioni meccaniche, senza però irrigidirsi in definizioni che la rapida evoluzione tecnologica renderebbe presto obsolete. La sfida consiste nel preservare la funzione selettiva dell'originalità – il che significa evitare che qualsiasi *output* computazionale riceva automaticamente protezione – senza tuttavia sacrificare la possibilità di riconoscere valore a creazioni autenticamente innovative anche nel momento in cui esse emergano da processi non umani (o non integralmente e completamente tali).

### 3. Il problema dell'intenzionalità nelle creazioni generate dall'IA

Oltre alle questioni dell'*agency* e dell'originalità, l'intelligenza artificiale generativa solleva un terzo problema filosofico, che attiene all'intenzionalità delle creazioni che essa produce. Per intenzionalità si intende – seguendo la tradizione filosofica inaugurata da Franz Brentano<sup>8</sup> – la proprietà degli stati mentali di essere diretti verso oggetti, di avere un contenuto rappresentazionale. Credere, desiderare, sperare, temere sono tutti stati intenzionali perché si riferiscono a qualcosa: “si crede” che piova, “si desidera” un gelato, “si spera” nel successo, “si teme” il fallimento. Il quesito circa la possibilità o meno dell'IA di avere stati intenzionali è cruciale per il diritto d'autore perché la creazione artistica è tradizionalmente intesa come manifestazione dell'intenzionalità dell'autore. Altrimenti detto, l'artista “vuole esprimere” determinate idee, “vuole” evocare certe emozioni, “vuole” comunicare alcuni significati specifici.

Se l'IA non possiede intenzionalità – ovverosia se i suoi *output* non esprimono alcuna intenzione comunicativa o espressiva – allora le opere generate algoritmicamente differiscono ontologicamente da quelle umane,

<sup>8</sup> F. BRENTANO, *Sull'origine della conoscenza morale*, trad. it., Brescia, Scholé, 2025.

anche nel momento in cui, a livello fattuale, siano in tutto e per tutto indistinguibili. Questa differenza ontologica potrebbe giustificare un trattamento giuridico diversificato. Tuttavia, come si vedrà, la questione dell'intenzionalità artificiale è filosoficamente controversa. Non a caso, vi sono varie posizioni teoriche in merito, che conducono a conclusioni giuridiche assai divergenti. Di qui, la necessità di chiarire la relazione tra intenzionalità e creatività muovendo dalla ripresa – seppur sintetica – di tre diverse prospettive, vale a dire: quella della filosofia della mente contemporanea; quella della rilevanza giuridica dell'assenza di coscienza fenomenica; e, infine, quella dell'intenzionalità mediata, cioè dell'idea che le opere generate dall'IA possano veicolare l'intenzionalità dell'utilizzatore umano.

L'idea che i sistemi artificiali possano possedere una qualche forma di intenzionalità, negli ultimi cinquant'anni, è stata al centro del dibattito degli studiosi della filosofia della mente. Fra tutti, il già citato Franz Brentano<sup>9</sup> aveva individuato nell'intenzionalità il criterio distintivo fra ciò che è “mentale” e ciò che, invece, è “fisico”. In quest'ottica, solo i fenomeni psichici manifestano la proprietà di riferirsi a oggetti che possono anche non esistere (si pensi alla possibilità di desiderare un unicorno). Da par sua, Edmund Husserl sviluppò ulteriormente questa stessa intuizione nell'ambito della fenomenologia, distinguendo tra l'atto intenzionale (*noesis*) e l'oggetto inteso (*noema*)<sup>10</sup>. Speculazioni che sono proseguite poi nel ventesimo secolo, con la filosofia analitica, che si è interrogata sulla possibilità che l'intenzionalità sia riducibile a proprietà fisiche o funzionali, o se, piuttosto, essa rappresenti un aspetto irriducibile della coscienza.

L'avvento dell'IA – e non poteva essere altrimenti – ha ulteriormente radicalizzato il dibattito. Alcuni filosofi, come Hilary Putnam e Jerry Fodor, hanno sostenuto una concezione funzionalista dell'intenzionalità: uno stato mentale è definito non dalla sua costituzione materiale ma dal ruolo causale che svolge all'interno del sistema cognitivo complessivo. In questa prospettiva, se un sistema computazionale implementa gli stessi ruoli funzionali di uno stato intenzionale umano, allora possiede intenzionalità genuina. Un computer che gioca a scacchi “crede” che il re sia in pericolo e “desidera” proteggere i propri pezzi nello stesso senso funzionale in cui lo crede e desidera un giocatore umano.

Questa posizione è stata vigorosamente contestata da John Searle attraverso l'esperimento mentale della *Stanza cinese*. Searle immagina una

<sup>9</sup> F. BRENTANO, *Psicologia dal punto di vista empirico*, trad. it., Bari, Laterza, 1997.

<sup>10</sup> E. HUSSERL, *Idee per una fenomenologia pura e per una filosofia fenomenologica*, trad. it., Torino, Einaudi, 2002.

persona che non comprende il cinese chiusa in una stanza con un manuale di regole sintattiche: l'individuo riceve domande scritte in cinese attraverso una fessura, consulta il manuale per elaborare risposte appropriate e le restituisce attraverso un'altra fessura. Dall'esterno, il sistema sembra comprendere il cinese, ma, nella realtà dei fatti, la persona nella stanza non comprende nulla, si limita a manipolare simboli secondo regole formali. Analogamente, argomenta Searle, un computer può manipolare simboli secondo algoritmi pur senza comprendere il loro significato, dunque, pur senza possedere un'autentica intenzionalità.

L'argomento di Searle – alquanto noto oltre che straordinariamente esemplare – ha dato il via a riflessioni ulteriori e a risposte diverse e decisamente articolate. Da par suo, la *reply* dei sistemi sostiene che – anche se il singolo individuo non comprende – il sistema nel suo complesso (individuato dall'*ensemble*: persona+manuale+stanza) comprende il cinese. La *reply* del robot, invece, argomenta che l'intenzionalità richiederebbe l'incorporazione (*embodiment*, sulla critica di Dreyfus all'IA come sistema privo di radicamento corporeo) del sistema in un corpo capace di interagire causalmente con l'ambiente, ma non è impossibile in linea di principio. Daniel Dennett sul punto ha proposto una via per così dire intermedia, vale a dire quella di distinguere tra l'intenzionalità originale o intrinseca, che caratterizza gli esseri umani, e l'intenzionalità derivata, che può essere attribuita ai sistemi artificiali dall'interprete esterno. Se così, un termostato “crede” che la temperatura sia al di sotto del valore impostato e “desidera” riscaldare l'ambiente, ma solo in senso derivato.

L'IA generativa complica ulteriormente il quadro. I modelli *transformer* che producono testi, ad esempio, non manipolano simboli discreti secondo regole esplicite, ma operano attraverso reti neurali che apprendono rappresentazioni distribuite. Quando *GPT* genera/formula la frase “Il gatto è sul tappeto”, il sistema denota una rappresentazione interna che corrisponde al significato della frase? O, piuttosto, sta semplicemente predicendo le sequenze di *token* più probabili? Alcuni ricercatori in *AI interpretability* sostengono che i modelli linguistici sviluppano rappresentazioni semantiche genuine durante l'addestramento; altri, invece, ritengono che si tratti di mera correlazione statistica senza comprensione effettiva.

Dal punto di vista del diritto d'autore, questa controversia di stampo filosofico determina una serie di conseguenze dirette. Se l'IA possiede intenzionalità genuina – anche se derivata o di grado inferiore a quella umana – allora i suoi *output* potrebbero essere considerati espressioni autentiche e, dunque, meritevoli di tutela. Se, all'opposto, l'IA è priva di intenzionalità, allora le sue produzioni sono semplicemente dei fenomeni

meccanici privi di quel carattere espressivo che giustifica la protezione autorale. Vien da sé che la difficoltà risiede nel fatto che l'intenzionalità non è direttamente osservabile dall'esterno: possiamo, infatti, solo tentare di inferirla a partire dal comportamento del sistema. Il che, ovviamente, significa che se il comportamento è indistinguibile da quello intenzionale, la distinzione rischia di diventare sfuggente, prettamente metafisica e – quel che più conta in questa sede – concretamente e giuridicamente inservibile e inapplicabile.

Un aspetto particolare del problema dell'intenzionalità riguarda la coscienza fenomenica o qualitativa, vale a dire, il carattere soggettivo dell'esperienza, ossia “che cosa si prova” a vedere il rosso, “che cosa si prova” ad ascoltare musica, “che cosa si prova” a provare emozioni. Anche ammettendo che l'IA possa possedere intenzionalità in senso funzionale, rimane la questione se essa possa avere esperienze qualitative. La rilevanza giuridica che tale distinzione ha con riguardo al diritto d'autore richiede che si valuti se la tutela debba (o meno) presupporre il vissuto interiore del processo creativo da parte dell'autore. In alternativa, occorre valutare se sia sufficiente che il risultato finale presenti quelle caratteristiche estetiche che sono apprezzabili e sufficienti per giustificare l'inclusione nella protezione normativa.

La filosofia della mente distingue tra quello che è il problema “facile” della coscienza (cioè, spiegare le funzioni cognitive come attenzione, memoria, elaborazione dell'informazione) e quello che è il problema “difficile” (ovvero, spiegare perché queste funzioni sono accompagnate da esperienza soggettiva). David Chalmers ha sostenuto che nessuna spiegazione puramente funzionale, o prettamente fisica, può cogliere l'aspetto qualitativo della coscienza. In tal senso, a mo' di paradosso, si potrebbe immaginare perfino una sorta di uno “zombie filosofico”<sup>11</sup>, in tutto e per tutto identico a un essere umano – almeno sotto il profilo funzionale – ma assolutamente privo di esperienza soggettiva. Seguendo questa lettura, allora anche un'IA che sia perfettamente equivalente a un essere umano sul piano computazionale potrebbe difettare di coscienza fenomenica.

Thomas Nagel, da par sua, nel celebre saggio *Che cosa si prova a essere un pipistrello?*<sup>12</sup>, sostiene che la coscienza è essenzialmente soggettiva e non può essere colta a partire da prospettive oggettive “di terza perso-

<sup>11</sup> D.J. CHALMERS, *The Conscious Mind: In Search of a Fundamental Theory*, Oxford, Oxford University Press, 1996.

<sup>12</sup> T. NAGEL, “What Is It Like to Be a Bat?”, in *The Philosophical Review*, 83(4), 1974, pp. 435-450.

na”. In questo senso, l’esperienza del pipistrello che percepisce il mondo attraverso l’ecolocalizzazione ci è – e ci resta – radicalmente inaccessibile. Analogamente, se l’IA avesse una forma di coscienza, essa sarebbe così diversa dalla nostra da risultare comunque sia inconoscibile. Ma più probabilmente – prosegue Nagel – i sistemi computazionali sono privi di qualsiasi esperienza soggettiva perché sono privi di quella particolare struttura biologica che, negli organismi viventi, dà origine alla coscienza.

Sulla scorta di queste considerazioni è possibile chiedersi se l’assenza di coscienza fenomenica nell’IA sia giuridicamente rilevante per il diritto d’autore. Una prima possibile risposta a questo quesito è negativa: questo perché ciò che conterebbe giuridicamente non sarebbe l’esperienza interiore dell’autore bensì il risultato oggettivo dell’opera. Altrimenti detto, nel momento in cui un’opera possiede qualità estetiche rilevanti, originalità formale, capacità comunicativa, allora essa risulta meritevole di tutela indipendentemente dal fatto che il suo creatore abbia vissuto interiormente il processo creativo. Questa posizione ha il vantaggio della praticabilità. Del resto, proprio perché la coscienza è un fenomeno interno e non verificabile dall’esterno, basare la tutela giuridica su tale criterio rischierebbe di rendere il diritto stesso assolutamente arbitrario.

Una seconda risposta è invece affermativa: la protezione del diritto d’autore non mirerebbe solo a incentivare la produzione di opere culturali, ma anche a riconoscere e a tutelare la dignità creativa del suo ideatore. La creatività in questo senso è percepita come espressione della soggettività, della capacità di immaginare mondi possibili, di provare emozioni e di trasmetterle attraverso forme simboliche. Se così, è evidente che un’opera generata da un sistema privo di coscienza – per quanto esteticamente apprezzabile – non è che espressione di un soggetto e quindi non può ricevere la stessa tutela riservata alle creazioni umane. Questa posizione riflette una concezione personalista del diritto d’autore, coerente con la tradizione giuridica continentale e con le teorie hegeliane che in questa sede sono già state analizzate.

La posizione intermedia elaborata da Luciano Floridi – e accennata in precedenza – conduce a una conclusione normativa assolutamente coerente con questo quadro. Nel dettaglio, le opere generate dall’IA potrebbero ricevere una tutela proporzionale al grado di intenzionalità coinvolto nel processo creativo, protezione piena per opere create da esseri umani consapevoli, protezione ridotta o *sui generis* per quelle generate da sistemi con intenzionalità funzionale ma non riflessiva. Ciò che importa sottolineare in questo contesto specifico è che tale gradazione non è puramente filosofica: corrisponde, piuttosto, a una scelta norma-

tiva ben precisa, che è coerente con la posizione personalista del diritto d'autore continentale, secondo cui la tutela piena è riservata solo – e soltanto – a manifestazioni di soggettività consapevole.

In questa disamina si abbraccia questa terza posizione, riconoscendo che la coscienza fenomenica – pur non essendo un requisito tecnico della tutela autorale – è comunque un valore che il diritto deve considerare. La creatività umana – in quanto manifestazione di un soggetto consapevole – merita un riconoscimento privilegiato che non può essere esteso automaticamente a produzioni meccaniche, anche quando esse siano esteticamente sofisticate. Nonostante ciò, però, questo privilegio non deve tradursi in un diniego assoluto di ogni tutela per le opere generate dall'IA, quanto piuttosto in un bilanciamento attento e ponderato degli interessi coinvolti e in forme differenziate di protezione.

Una terza prospettiva circa il problema dell'intenzionalità delle opere generate dall'IA è quella dell'intenzionalità “mediata”. Si tratta dell'idea secondo la quale, anche se il sistema artificiale non possiede intenzionalità propria, esso è comunque in grado di veicolare l'intenzionalità dell'essere umano che lo utilizza. Ad esempio, quando un artista usa *Midjourney* per generare un'immagine descrivendo dettagliatamente lo stile desiderato, i colori, così come la composizione, e itera molteplici volte modificando i *prompt* fino a ottenere il risultato voluto, in che senso l'*output* finale può essere considerato espressione dell'intenzionalità dell'artista piuttosto che mera produzione automatica del sistema?

Questa questione non è nuova: si pone ogni volta che uno strumento tecnologico media tra l'intenzione dell'autore e l'opera finale. Il fotografo non dipinge direttamente l'immagine ma configura parametri tecnici (inquadratura, esposizione, messa a fuoco) mentre – tecnicamente e fattualmente – è il processo chimico (o digitale) che produce l'immagine. Il compositore che usa un sintetizzatore non genera direttamente le onde sonore ma programma oscillatori, filtri, modulatori che producono il suono finale. Ciò che rende queste attività creative non è l'esecuzione materiale ma la serie di scelte intenzionali che determinano il risultato. Seguendo l'esempio, la fotografia è giuridicamente proteggibile non perché il fotografo abbia fisicamente dipinto l'immagine, ma perché egli ha compiuto scelte creative relative al soggetto, all'inquadratura, all'illuminazione.

L'IA generativa può essere proficuamente inquadrata in questo paradigma della c.d. creatività mediata tecnologicamente. Qui, invero, l'utilizzatore fornisce istruzioni (*prompt*), configura parametri (modello, risoluzione, stile), seleziona tra *output* alternativi e, eventualmente, modifica il risultato. Ciascuna di queste fasi incorpora scelte intenzionali

che riflettono la visione creativa dell'utilizzatore. In questa prospettiva, l'IA è uno strumento sofisticato che amplifica le capacità espressive umane, analogo alla fotocamera o al sintetizzatore. L'intenzionalità dell'opera risiede, dunque, non già nel sistema bensì nell'utilizzatore che lo dirige verso specifici obiettivi creativi.

Chiaramente, il ricorso a questa analogia incontra difficoltà quando il contributo umano è minimo. Se l'utilizzatore inserisce un *prompt* generico (genera un paesaggio) e accetta il primo *output* senza modifiche, è difficile sostenere che l'immagine risultante esprima una visione creativa personale. Il grado di intenzionalità veicolato sembra inversamente proporzionale all'autonomia del sistema: più il sistema è capace di produrre risultati sofisticati con *input* minimali, meno l'utilizzatore può rivendicare che l'*output* finale rifletta le sue scelte creative. Questo suggerisce la necessità di criteri graduati che valutino caso per caso il contributo intenzionale dell'utilizzatore.

Una proposta interessante – avanzata da Pamela Samuelson – è quella di distinguere tra diversi tipi di utilizzo dell'IA in base al grado di controllo creativo esercitato dall'utilizzatore. Un primo livello è quello dell'uso strumentale: l'IA assiste l'autore umano eseguendo compiti specifici all'interno di un progetto creativo più ampio (ad esempio, generare uno sfondo per un'illustrazione). Un secondo livello è quello dell'uso collaborativo: l'utilizzatore e il sistema interagiscono iterativamente, ciascuno contribuendo al risultato finale in modi complementari. Un terzo livello è quello dell'uso generativo puro: il sistema produce autonomamente l'intera opera con intervento umano minimo. E di questi, solo i primi due livelli giustificherebbero l'attribuzione della paternità all'utilizzatore umano.

Un problema ulteriore riguarda la trasparenza del processo generativo. Quando l'utilizzatore fornisce un *prompt* a un modello di IA, non sa esattamente come il sistema elaborerà le istruzioni né quali elementi del *dataset* di addestramento influenzeranno l'*output*. Questa opacità rende difficile stabilire in che misura l'*output* rifletta effettivamente l'intenzionalità dell'utilizzatore o incorpori invece *pattern* appresi autonomamente dal modello. Se il sistema è stato addestrato su opere di uno specifico artista e riproduce involontariamente lo stile di quell'artista, l'utilizzatore può rivendicare che l'opera esprima la sua intenzionalità creativa?

L'analisi si propone di affrontare questa questione adottando un criterio pragmatico: l'intenzionalità mediata è riconosciuta giuridicamente quando l'utilizzatore può dimostrare di aver esercitato un controllo creativo significativo sul processo generativo. Questo controllo può manifestarsi attraverso: la specificità e complessità dei *prompt*; l'iterazione e raffinamento degli *output*; la selezione consapevole tra alternative; l'integrazione dell'*output*

in un contesto creativo più ampio; la modificazione e la personalizzazione del risultato. Quanto maggiore è il controllo esercitato, tanto più forte è l'argomentazione che l'opera finale sia il risultato dell'intenzionalità dell'utilizzatore umano e pertanto meriti protezione autorale.

Questa soluzione non risolve tutti i problemi teorici – in particolare, non fornisce un parametro e una definizione metafisica precisa di quando un'opera “esprime” intenzionalità – ma offre un criterio operativo che i tribunali possono applicare valutando le circostanze concrete di ciascun caso. In pratica, si privilegia la flessibilità e il contestualismo rispetto a definizioni generali e rigide, che, fra le altre cose, la rapidità dell'evoluzione tecnologica renderebbe presto obsolete e inservibili allo scopo.

#### 4. *La dimensione temporale della creatività artificiale*

Nei tre paragrafi precedenti si è considerata quella che è la sfida ontologica che l'IA generativa pone al diritto d'autore, concentrandosi in modo particolare sulle questioni di *agency*, originalità e intenzionalità. A questo punto è necessario rivolgere l'attenzione a un aspetto ulteriore – spesso trascurato ma giuridicamente cruciale – ovvero alla dimensione per così dire temporale della creatività artificiale. L'IA generativa, infatti, non solo produce opere in modi ontologicamente molto diversi rispetto alla creatività umana, ma le produce anche con una velocità e con una frequenza radicalmente diverse. Questa differenza quantitativa determina conseguenze qualitative sul sistema del diritto d'autore, mettendo in discussione assunti impliciti circa i tempi della creazione, la scarsità relativa delle opere e la durata appropriata della protezione.

Si badi: la creatività umana è intrinsecamente vincolata dalla temporalità biologica. Scrivere un romanzo richiede mesi o anni; comporre una sinfonia richiede settimane; dipingere un quadro richiede giorni. Anche nelle forme creative più rapide – l'improvvisazione musicale, la poesia estemporanea – vi sono limiti fisici e cognitivi alla velocità di produzione. Questi vincoli temporali hanno influito storicamente sull'elaborazione e sul sistema di tutela del diritto d'autore: la durata della protezione (originariamente quattordici anni, poi estesa progressivamente per tutta la durata della vita dell'autore e sino a settanta anni dopo la sua morte) presuppone che la creazione richieda investimenti significativi di tempo e che l'autore meriti un periodo prolungato per sfruttare economicamente l'opera ideata. La scarsità relativa delle opere creative giustifica l'attribuzione di diritti esclusivi: se la produzione culturale fosse sovrabbondante, i monopoli temporanei sarebbero meno necessari.

L'IA generativa sovverte pressoché tutti questi assunti. Un sistema come *DALL-E* può generare centinaia di immagini in pochi secondi; *GPT* può produrre un romanzo completo in alcuni minuti; sistemi di composizione musicale possono generare migliaia di brani in un'ora.

La velocità computazionale dell'IA trasforma la creatività da attività lenta e laboriosa a processo istantaneo e seriale. Le conseguenze giuridiche di questa metamorfosi sono molteplici e impongono al diritto di definire nuovi criteri per la regolamentazione della produzione massiva. In particolare, risulta problematico giustificare l'estensione di una protezione pluridecennale a opere generate in frazioni di secondo. Emerge, inoltre, l'esigenza di elaborare meccanismi preventivi per evitare che la saturazione dello spazio culturale da parte della creatività algoritmica renda impossibile, o economicamente insostenibile, la sopravvivenza della creatività umana tradizionale. Per saturazione dello spazio culturale si intende qui il fenomeno per cui la produzione massiva e seriale di *output* algoritmici – testi, immagini, composizioni musicali – occupa progressivamente quegli spazi di visibilità, di distribuzione e di remunerazione sui quali si regge l'ecosistema economico della creatività umana. Quando un sistema di IA è in grado di generare in pochi minuti più contenuti di quanti un autore umano ne produca in un'intera carriera, il mercato rischia di premiare la quantità e l'accessibilità a discapito della qualità e dell'autenticità. Il rischio non è astratto, i primi studi empirici segnalano già effetti di spiazzamento e di contrazione della domanda di lavoro creativo umano in alcuni segmenti del mercato, come l'illustrazione, la produzione di contenuti testuali standardizzati e altre attività creative ripetitive, nei quali i sistemi generativi offrono alternative a costo marginale tendente a zero. In quest'ottica, il problema non è soltanto di diritto della proprietà intellettuale, ma investe più ampiamente le condizioni materiali di esercizio della creatività umana come pratica culturale e professionale. Una regolamentazione adeguata dovrebbe quindi prevedere, accanto ai tradizionali strumenti autoriali, meccanismi di compensazione e di riserva di mercato che tutelino la sostenibilità economica degli autori umani di fronte alla concorrenza algoritmica sistemica<sup>13</sup>.

<sup>13</sup> Sul rischio di marginalizzazione economica dei creativi umani nell'ecosistema dell'IA generativa si vedano, con specifico riferimento al mercato dell'illustrazione e del giornalismo: G. COMANDINI, "Arte, intelligenza artificiale e diritto d'autore: prime riflessioni", in *Rivista di diritto industriale*, 2023, pp. 3-28; P. SAMUELSON, "Generative AI Meets Copyright", in *Science*, 381(6654), 2023, pp. 1293-1294; nonché il rapporto empirico di E. JIANG, K. NAOUS, A. RAMBACHAN, "Artificial Intelligence and Jobs: The Role of Demand", *NBER Working Paper*, n. 31870, 2023. Per una prospettiva filosofico-giu-

Una delle prime conseguenze della diversa temporalità che contraddistingue le opere prodotte dall'IA riguarda il divario tra la velocità computazionale della creazione algoritmica e la temporalità insita nelle norme del diritto d'autore. Come ha osservato acutamente Lawrence Lessig<sup>14</sup>, il diritto è una forma di regolazione che presuppone determinate costanti fattuali. Ad esempio: la legislazione sul traffico presuppone che le automobili si muovano a velocità limitate; il diritto penale presuppone che gli esseri umani possano essere identificati e localizzati; il diritto d'autore presuppone che la creazione richieda tempo e sforzo. Ragion per cui, quando le costanti fattuali cambiano drasticamente, le norme tradizionali possono diventare inadeguate o addirittura controproducenti.

Nel caso del diritto d'autore, la velocità computazionale dell'IA sfida almeno tre assunti temporali impliciti.

Il primo riguarda il momento della creazione: tradizionalmente, un'opera viene creata attraverso un processo temporalmente esteso in cui l'autore elabora progressivamente idee, corregge bozze, affina dettagli. L'opera finale è il risultato di un'evoluzione temporale in cui scelte successive si stratificano. L'IA, invece, genera l'opera istantaneamente a partire dall'*input* iniziale: non vi è processo temporale di elaborazione ma una trasformazione computazionale immediata da *prompt* a *output*. Questa istantaneità solleva interrogativi circa l'applicabilità di categorie giuridiche che presuppongono la temporalità: che senso ha parlare di "versioni successive" o di "evoluzione creativa" quando l'opera emerge compiuta in microsecondi?

Il secondo assunto riguarda la relazione tra investimento temporale e valore economico. La teoria giuslavorista di Locke – analizzata in precedenza – giustifica la proprietà come ricompensa per il lavoro investito. Analogamente, la giustificazione utilitarista del *copyright* si basa sull'idea che – senza incentivi economici – gli autori non investirebbero tempo ed energie nella creazione. Ma se la creazione richiede tempo

ridica più ampia sul rapporto tra automazione tecnologica e condizioni di esercizio della creatività umana, si veda M. FASAN, *Creatività e intelligenza artificiale*, Napoli, Edizioni Scientifiche Italiane, 2023.

<sup>14</sup> L. LESSIG, *Code and Other Laws of Cyberspace*, New York, Basic Books, 1999. La tesi della dipendenza strutturale del diritto da costanti fattuali presupposte è sviluppata anche in L. LESSIG, *Free Culture: How Big Media Uses Technology and the Law to Lock Down Culture and Control Creativity*, New York, Penguin Press, 2004, dove l'autore osserva che ogni sistema normativo è calibrato su assunti circa la realtà che intende regolare, e che il mutamento tecnologico può rendere tali assunti obsoleti senza che le norme vengano formalmente modificate, generando inadeguatezza strutturale del diritto rispetto ai fenomeni che pretende di disciplinare.

computazionale trascurabile, l'incentivo economico è ancora necessario? Un'obiezione possibile è che, anche se la generazione della singola opera è istantanea, lo sviluppo del sistema di IA richiede investimenti enormi in ricerca, dati, infrastrutture computazionali. La protezione non remunera il tempo della generazione ma l'investimento nell'infrastruttura tecnologica. Ovviamente, questa risposta è problematica. Se si protegge l'opera generata dall'IA in virtù dell'investimento nell'infrastruttura, allora la titolarità dovrebbe spettare agli sviluppatori del sistema piuttosto che agli utilizzatori. Ma questo contrasta con l'idea, difesa da alcuni, che l'utilizzatore dovrebbe essere considerato autore. Inoltre, questa logica condurrebbe a monopolizzare vaste porzioni dello spazio culturale: se un'azienda sviluppa un sistema capace di generare milioni di opere, tutte protette per decenni, essa acquisisce un controllo senza precedenti sulla produzione simbolica. Come ha affermato James Boyle<sup>15</sup>, il rischio è una seconda recinzione dei *commons*, in cui conoscenze e forme espressive precedentemente accessibili vengono privatizzate attraverso la proprietà intellettuale estesa alle opere algoritmiche.

Il terzo assunto riguarda la durata della protezione. La tutela prolungata (calibrata sulla vita dell'autore ed estesa a settanta anni dalla sua morte) si giustifica in parte con l'idea che l'opera rappresenta l'eredità culturale dell'autore, destinata a sopravvivergli, e che i discendenti meritano di beneficiare economicamente del talento del loro progenitore. Ci si chiede, se, oggi, nel momento in cui l'IA è in condizioni di generare milioni di opere in pochi giorni, abbia ancora senso estendere una protezione così lunga. L'obsolescenza culturale delle opere algoritmiche potrebbe essere molto più rapida di quella delle opere umane: contenuti generati massivamente per rispondere a mode passeggere perdono rilevanza in tempi brevissimi. Proteggere tali contenuti per decenni significa invece andare a vincolare utilizzi futuri senza che si abbia un beneficio sociale proporzionato alla tutela approntata.

L'istantaneità della creazione algoritmica, inoltre, ha conseguenze normative che trascendono la questione della durata della protezione, andando a interessare anche la definizione stessa di opera e la struttura dei diritti esclusivi. Nel sistema tradizionale del diritto d'autore, la protezione sorge automaticamente con la creazione dell'opera, ossia con la sua fissazione in una forma tangibile. Questo automatismo presuppone che vi sia un momento preciso e identificabile in cui l'opera nasce, pas-

<sup>15</sup> J. BOYLE, *The Second Enclosure Movement and the Construction of the Public Domain*, in *Law and Contemporary Problems*, 66(1/2), 2003, pp. 33-74.

sando dal *non-essere*, all'*essere*. Per le opere umane, questo momento è relativamente chiaro: si identifica nel momento in cui l'autore termina la scrittura del manoscritto, completa il dipinto, fissa la registrazione audio.

Per le opere generate dall'IA, invece, il momento della creazione risulta decisamente più sfuggente e più sfumato. Consideriamo un sistema di IA addestrato su un *corpus* di opere: in un certo senso, tutte le opere potenziali che il sistema può generare esistono già implicitamente nello spazio latente delle rappresentazioni apprese. La generazione di un'opera specifica è semplicemente la selezione e la materializzazione di una delle infinite possibilità già codificate nel modello. In questa prospettiva, il sistema è come una biblioteca borgesiana<sup>16</sup> che contiene tutte le opere possibili: generare un'opera non significa più creare qualcosa di nuovo ma estrarre – dare forma e fissare – un qualcosa che, *in nuce*, è già esistente virtualmente.

Questa metafora ha conseguenze giuridiche rilevanti, ma non conduce necessariamente a un'*impasse*. La domanda che fa da cornice al ragionamento – ovvero: “quando sorge il diritto d'autore per un'opera algoritmica?” – chiama in causa tre possibili momenti: *i*) l'addestramento del modello; *ii*) la formulazione del *prompt*; *iii*) la materializzazione e fissazione dell'*output*. Le prime due opzioni sono strutturalmente inadeguate.

*i*) Se i diritti sorgessero nel momento dell'addestramento, lo sviluppatore sarebbe il titolare di tutte le infinite opere potenzialmente generabili dal sistema. Una soluzione, questa, che si rivela palesemente assurda non foss'altro perché trasformerebbe il diritto d'autore in un monopolio illimitato sull'intero spazio combinatorio del modello.

*ii*) Se i diritti sorgessero nel momento in cui si formula il *prompt*, occorrerebbe stabilire quando una sequenza di istruzioni possa essere considerata sufficientemente determinata per individuare un'opera spe-

<sup>16</sup> La metafora della biblioteca borgesiana rinvia al celebre racconto di J.L. BORGES, “*La biblioteca de Babel*” (1941), in Id., *Finzioni*, Torino, Einaudi, 1955. Il racconto descrive una biblioteca che contiene tutti i libri possibili – ossia tutte le combinazioni ammissibili di caratteri entro un formato dato – e dunque anche tutti i libri che potranno mai essere scritti. La biblioteca borgesiana non è il risultato di un atto creativo, ma la materializzazione esaustiva di uno spazio combinatorio: in essa non esiste scoperta, ma soltanto estrazione di ciò che è già implicitamente contenuto. L'analogia con lo spazio latente dei modelli generativi è filosoficamente suggestiva: anche il modello addestrato “contiene”, in senso statistico e distribuito, tutte le opere che potrà generare. Tuttavia, come si argomenta nel testo, tale analogia non esaurisce la questione giuridica: il diritto d'autore non protegge possibilità virtuali ma iscrizioni effettive, e il momento della fissazione dell'output su supporto rimane il momento costitutivo rilevante.

cifica: criterio inevitabilmente arbitrario, data l'*eterogeneity* dei *prompt* e la non linearità tra *input* e *output*.

iii) La terza opzione – vale a dire la materializzazione e la fissazione dell'*output* su supporto – è l'unica coerente con i fondamenti teorici dell'indagine qui condotta. Come stabilito nella premessa metodologica, la teoria della documentalità di Ferraris<sup>17</sup> individua nella formula "*oggetto sociale = atto iscritto*" la struttura ontologica costitutiva di ogni oggetto giuridicamente rilevante. Un fatto diventa oggetto sociale nel momento in cui non esiste come possibilità virtuale, ma quando viene fissato – letteralmente quando viene iscritto – su un supporto che lo rende stabile, accessibile e socialmente percepibile. Prima della fissazione, l'*output* algoritmico è un fatto bruto potenziale; con la fissazione, diventa un fatto istituzionale su cui il diritto può operare. Questa prospettiva conduce a una conclusione netta: il diritto sorge con la fissazione dell'*output* su supporto e non in alcun momento anteriore. La metafora borgesiana descrive correttamente lo spazio latente del modello, ma quello spazio è – almeno nella terminologia ferrarisiana – un insieme di fatti "bruti", potenziali privi di rilevanza giuridica. E, come sappiamo, il diritto non protegge possibilità, protegge iscrizioni.

L'obiezione più seria a questa soluzione riguarda gli *output* identici generati indipendentemente da utilizzatori diversi: se entrambi fissassero lo stesso *output*, entrambi avrebbero protezioni concorrenti sulla stessa immagine o sullo stesso testo. Questo problema non è tuttavia esclusivo delle opere algoritmiche: il diritto d'autore ha già affrontato situazioni analoghe attraverso l'*independent creation doctrine*, stabilendo che due opere identiche sono entrambe protette purché siano state create autonomamente, senza copiatura. Nel contesto algoritmico la soluzione è analoga: due *output* identici generati da *prompt* indipendenti da utilizzatori diversi sono entrambi protetti, ma nessuno dei due vanta un diritto esclusivo sull'altro. La coincidenza non configura violazione perché manca la copia. Questa soluzione richiede però un sistema di documentazione del processo generativo – data, *prompt*, configurazioni del modello – che permetta di verificare l'indipendenza della creazione. Tale requisito di tracciabilità è coerente con il principio di trasparenza e *accountability*, che, nella premessa metodologica, è già stato individuato come uno degli elementi distintivi e imprescindibili di questo lavoro.

La soluzione fondata sulla fissazione non risolve, da sola, tutti i problemi temporali posti dalla creatività algoritmica. Rimane aperta la que-

<sup>17</sup> M. FERRARIS, *Documentalità. Perché è necessario lasciar tracce*, Roma-Bari, Laterza, 2009.

stione delle opere che esistono in uno stato di variazione continua senza una forma canonica definitiva. Si tratta di quegli *output* transienti di processi generativi continui che non vengono mai fissati in modo stabile. È su questo residuo, e non sulla questione principale del momento costitutivo, che si concentra la proposta di Guadamuz<sup>18</sup>.

Un secondo problema riguarda la modificabilità delle opere algoritmiche. Tradizionalmente, modificare un'opera esistente può dar luogo a un'opera derivata, protetta indipendentemente purché incorpori un contributo creativo sufficiente. Ma nel caso dell'IA generativa, la distinzione tra opera originale e opera derivata si perde. Il processo di generazione iterativa, in cui l'utilizzatore modifica progressivamente il *prompt* per ottenere variazioni dell'immagine, pone il problema di stabilire se ogni iterazione costituisca un'opera autonoma o se l'intero ciclo debba essere ricondotto a un unico processo creativo unitario. Allo stesso modo, la divergenza creativa tra due utilizzatori che partono dal medesimo *prompt* per apportare variazioni distinte richiede di definire se le opere risultanti debbano essere considerate creazioni indipendenti o se sussista un rapporto di derivazione tra l'una e l'altra. La soluzione fondata sulla fissazione risolve la questione principale del momento costitutivo, ma lascia aperto un problema residuale specifico delle opere algoritmiche: quello degli *output* che esistono in uno stato di variazione continua senza mai essere fissati in forma stabile. Per questa categoria – che Andres Guadamuz propone di denominare “opere fluide” – la tutela tradizionale basata sulla fissazione risulta strutturalmente inadeguata. Nel dettaglio, per le opere fluide, la tutela tradizionale basata sulla fissazione e sulla permanenza si rivela inadeguata. Occorrerebbe, invece, una forma di protezione più flessibile, forse basata su parametri generativi (*prompt*, *seed*, configurazioni) piuttosto che su *output* specifici. Ovviamente, questa soluzione solleva a sua volta problemi: proteggere i parametri generativi significa proteggere processi piuttosto che opere, avvicinandosi pericolosamente alla protezione delle idee vietata dal principio idea-espressione.

La terza dimensione del problema temporale riguarda l'obsolescenza culturale accelerata delle opere generate dall'IA e le sue implicazioni per la durata della protezione stessa. Il diritto d'autore attribuisce ai titolari diritti esclusivi per periodi molto lunghi: nell'Unione Europea, per la durata della vita dell'autore e per i settant'anni successivi alla sua morte;

<sup>18</sup> A. GUADAMUZ, “Do Androids Dream of Electric Copyright? Comparative Analysis of Originality in Artificial Intelligence Generated Works”, in *Intellectual Property Quarterly*, 2, 2017, pp. 169-186.

negli Stati Uniti, novantacinque anni per le opere corporative. Queste durate sono state progressivamente estese nel corso del ventesimo secolo, nonostante le critiche di coloro che denunciano l'eccessiva lunghezza dei monopoli e il conseguente impoverimento del dominio pubblico. Le giustificazioni tradizionali per durate così lunghe sono molteplici. Una giustificazione economica sostiene che la protezione estesa massimizza l'incentivo alla creazione. Altrimenti detto, sapendo che l'opera sarà protetta per molto tempo, l'autore sarà motivato a investire maggiormente nella sua realizzazione.

Un romanzo classico continua a essere letto decenni dopo la morte dell'autore; una composizione musicale può rimanere nel repertorio per secoli; un'opera d'arte può acquisire valore crescente nel tempo. Per tutte queste opere, una tutela temporalmente lunga ha senso: si tratta, infatti, di proteggere investimenti culturali duraturi. Ma molte opere contemporanee – specialmente nei media digitali – hanno cicli di vita molto più brevi: un videogioco può diventare obsoleto in pochissimi anni; un articolo online perde rilevanza in alcune settimane; un meme virale è dimenticato in pochi giorni.

Altrimenti detto ed in breve, le opere generate dall'IA tendono a mostrare un'obsolescenza ancora più rapida. In particolare, dal momento che possono essere prodotte massivamente e a basso costo, esse saturano rapidamente i mercati culturali e si sostituiscono reciprocamente a ritmi vertiginosi. Un'immagine generata per illustrare un articolo giornalistico perde ogni valore dopo la pubblicazione dell'articolo stesso; un testo generato per rispondere a una richiesta specifica diventa irrilevante una volta soddisfatta la richiesta. Proteggere tali contenuti "costitutivamente effimeri" per settant'anni dopo la morte dell'autore (o novantacinque anni nel caso di opere corporative) appare decisamente sproporzionato oltre che giuridicamente ed economicamente inefficiente e indesiderabile.

Alcuni studiosi hanno proposto di ridurre drasticamente la durata della protezione per le opere generate dall'IA. Pamela Samuelson, da par sua, suggerisce un termine di cinque o dieci anni, sufficiente per permettere lo sfruttamento commerciale iniziale ma abbastanza breve da evitare il blocco del dominio pubblico. Altri propongono forme di tutela differenziata: durata piena per opere che incorporano creatività umana significativa, durata ridotta per opere generate prevalentemente dall'IA. Questa soluzione, coerente con l'approccio graduato, richiede però criteri chiari atti a distinguere le diverse categorie, evitando incertezze e contenziosi.

Un problema ulteriore riguarda la quantità assoluta di opere protette. Anche se la durata fosse ridotta, la produzione massiva di opere algorit-

miche potrebbe comunque saturare lo spazio culturale e creare grovigli di autorizzazioni (*permission thicket*s), che rendono impossibile la creatività derivativa. Immaginiamo un sistema di IA che generi milioni di immagini, ciascuna protetta per dieci anni. Un artista umano che voglia creare un'opera incorporando elementi visivi esistenti si troverebbe di fronte a un numero ingestibile di diritti da negoziare, anche se la durata di ciascuno è relativamente breve. Questo fenomeno di inquinamento del dominio pubblico richiederebbe soluzioni complementari: licenze automatiche, eccezioni ampliate, sistemi di compensazione collettiva.

La questione della durata si intreccia con quella dell'identificazione delle opere. Nel sistema tradizionale, ogni opera è identificata univocamente e la sua protezione può essere verificata attraverso registri o attraverso l'analisi dell'opera stessa. Ma se un sistema di IA genera milioni di *output* effimeri e transienti, molti dei quali mai salvati o conservati, come si può determinare se un'opera successiva viola i diritti su un'opera generata precedentemente? La difficoltà pratica di tracciare e documentare la produzione algoritmica massiva potrebbe rendere *de facto* inapplicabile qualsiasi forma di tutela, indipendentemente dalla durata formale.

Il testo si propone di affrontare questa costellazione di problemi attraverso un ripensamento complessivo della dimensione temporale del diritto d'autore nell'era algoritmica. Piuttosto che estendere automaticamente alle opere generate dall'IA le durate previste per le opere umane, occorre elaborare criteri flessibili che calibrino la protezione in base a fattori quali: l'investimento creativo umano coinvolto; la rilevanza culturale prevista dell'opera; l'impatto sulla disponibilità di materiali per la creatività derivativa; la praticabilità amministrativa della tutela. Questo approccio, pur richiedendo valutazioni caso per caso più complesse, permette di bilanciare meglio gli interessi in gioco ed evitare tanto l'iperprotezione quanto la sotto-tutela.

## CAPITOLO IV

### SOGGETTIVITÀ GIURIDICA E IA

SOMMARIO: 1. Teorie della personalità giuridica applicabili all'IA. – 2. L'attribuzione dei diritti nell'ecosistema dell'IA generativa. – 3. Diritti fondamentali e IA.

#### 1. *Teorie della personalità giuridica applicabili all'IA*

La questione della soggettività giuridica dell'IA costituisce uno dei nodi teorici più complessi e controversi del dibattito contemporaneo. Come si è visto nei capitoli precedenti, il diritto d'autore è costruito attorno alla figura dell'autore-persona, soggetto dotato di personalità giuridica e capacità creativa. L'IA generativa – producendo opere che esibiscono caratteristiche formalmente analoghe a quelle delle creazioni umane – solleva la questione se debba essere riconosciuta come soggetto di diritto, capace di essere titolare di diritti e obblighi. Tale questione non è puramente di stampo teorico, ma implica conseguenze pratiche immediate: se l'IA può essere autrice, allora le opere da essa generate riceverebbero protezione giuridica; se, invece, essa non può esserlo, occorre individuare a chi attribuire la titolarità dei diritti o valutare se tali opere debbano rimanere nel dominio pubblico.

Il presente paragrafo esamina le diverse teorie della personalità giuridica e la loro applicabilità all'IA. Si tratterà, innanzitutto, di ricostruire la natura della persona giuridica quale finzione normativa, evidenziando come il diritto abbia storicamente riconosciuto soggettività a entità diverse dagli esseri umani. In secondo luogo, si analizzerà il dibattito contemporaneo sulla *electronic personhood* e sulla *AI personhood*, valutandone tanto la plausibilità quanto i possibili rischi. Infine, si esamineranno modelli alternativi che permettano di regolare la responsabilità per le attività dell'IA senza necessariamente attribuire ad essa una personalità giuridica.

Come si vedrà, la questione della soggettività giuridica dell'IA non ammette risposte univoche e definitive. Essa dipende da scelte normative che riflettono una differente sensibilità circa i fini del diritto, come pure, in ordine alle posizioni e agli interessi che esso deve tutelare. L'indagine qui condotta si propone di elaborare criteri che permettano di affrontare questa complessità senza cadere, né in un antropocentrismo rigido che rifiuti a priori qualsiasi forma di riconoscimento giuridico dell'IA, né in

un entusiasmo tecnologico acritico che equipari l'IA alle persone fisiche misconoscendone la diversità ontologica.

È importante sottolineare, sin da subito, che la personalità giuridica – lungi dall'essere un dato naturale – è piuttosto una qualificazione normativa attraverso la quale l'ordinamento decide chi riconoscere come soggetto di diritto, capace di essere titolare di situazioni giuridiche soggettive. Questa constatazione – per quanto sia ovvia ai giuristi – merita comunque di essere sottolineata perché confuta l'illusione che esista un criterio ontologico oggettivo idoneo a determinare chi sia “persona” per il diritto. Da questo punto di vista, anzi, la storia del diritto dimostra invece che i diversi confini della soggettività giuridica di volta in volta sono stati tracciati e ritracciati in modi diversi a seconda dei contesti storici, delle ideologie dominanti e, non da ultimo, anche in ragione delle esigenze pratiche.

Nel diritto romano, ad esempio, non tutti gli esseri umani erano considerati persone: gli schiavi erano *res* (cose), oggetti sottoposti al regime di proprietà e del tutto privi di capacità giuridica. Le donne, i minori e gli stranieri poi avevano una capacità limitata. Al contempo, determinate entità collettive – in particolare le *civitates* e le corporazioni – erano riconosciute come soggetti in grado di possedere beni e di contrarre obbligazioni. Riprendendo e, in parte, sviluppando ulteriormente questa prospettiva, il diritto canonico medievale elaborò la teoria delle *personae morales*, attribuendo personalità giuridica alla Chiesa e ad altri enti ecclesiastici.

Con l'avvento della modernità la nozione di personalità giuridica venne estesa e, per così dire, universalizzata, di modo che ogni essere umano iniziò a essere considerato soggetto di diritto, indipendentemente dal suo *status* sociale, dalla sua razza o dal suo genere. Parallelamente, il diritto moderno ha elaborato anche la categoria delle persone giuridiche – società, fondazioni, enti pubblici – che, pur non essendo individui biologici, sono trattate dall'ordinamento come soggetti capaci di agire giuridicamente. La natura di tali entità è stata oggetto di varie dispute teoriche: per alcuni, esse sono finzioni (come sosteneva Von Savigny<sup>1</sup>), per altri, sono realtà sociali autonome (secondo la teoria della realtà di Gierke<sup>2</sup>), e, per altri ancora, esse sono semplicemente delle costruzioni normative utili (come argomentano i realisti). Ora, indipendentemente dalla risposta che si preferisca dare a questa particolare questione metafisica, è chiaro che l'attribuzione di una personalità giuridica alle società commerciali risponde alla necessità di soddisfare una serie di esigenze

<sup>1</sup> F.C. VON SAVIGNY, *System des heutigen römischen Rechts*, vol. 2, Berlin, Veit, 1840.

<sup>2</sup> O. VON GIERKE, *Das deutsche Genossenschaftsrecht*, vol. 1, Berlin, Weidmann, 1868.

pratiche, quali, ad esempio: facilitare l'organizzazione di imprese collettive, limitare la responsabilità dei soci, permettere la continuità giuridica oltre la vita degli individui.

Questa – seppur davvero concisa – ricostruzione storica dà conto di come la personalità giuridica abbia rappresentato – e rappresenti tuttora – uno strumento flessibile che l'ordinamento può adattare alle proprie esigenze. Invero, non vi è ragione di principio per escludere che entità non umane possano essere riconosciute come soggetti di diritto, se tale riconoscimento serve al soddisfacimento di scopi normativi legittimi.

Tuttavia, la storia insegna anche che l'estensione della soggettività giuridica comporta sempre conseguenze sistemiche profonde e richiede una riflessione attenta in ordine ai valori che si intendono tutelare. Se, infatti, per un verso, riconoscere le società come persone ha permesso lo sviluppo del capitalismo moderno, per un altro verso, tale scelta ha anche sollevato problemi di responsabilità e di *accountability*. Analogamente, riconoscere l'IA come soggetto giuridico potrebbe favorire l'innovazione tecnologica, ma potrebbe anche generare effetti distorsivi sul sistema dei diritti e delle responsabilità.

Il dibattito contemporaneo sulla soggettività giuridica dell'IA si è concentrato sulla proposta di istituire una *electronic personhood* o una *AI personhood*, ossia: una forma di personalità giuridica specificamente adattata alle caratteristiche dei sistemi artificiali intelligenti. Tale proposta è stata avanzata in diversi contesti, da quello normativo (basti ricordare che il Parlamento europeo, con la risoluzione del 16 febbraio 2017 recante raccomandazioni alla Commissione sulle norme di diritto civile sulla robotica, ha invitato a esplorare anche la possibilità, nel lungo periodo, di uno specifico status giuridico per i robot più sofisticati, fino alla prospettiva di *electronic persons*), a quello dottrinale (in cui vari studiosi hanno difeso questa soluzione sostenendo che si trattasse della risposta più consona a risolvere i problemi di responsabilità e di attribuzione di diritti suscitati dall'IA).

Gli argomenti a favore dell'*AI personhood* sono principalmente di natura funzionalista e pragmatica. Se l'IA è capace di agire autonomamente, di prendere decisioni non previste dai programmatori e di produrre effetti giuridicamente rilevanti, allora dovrebbe essere riconosciuta come soggetto di diritto affinché possa essere imputata di responsabilità e divenire in un certo qual senso anche titolare di diritti. Lawrence Solum – antesignano e pioniere di questa tesi – ha argomentato che sistemi sufficientemente sofisticati potrebbero soddisfare i requisiti minimi richiesti per il riconoscimento della personalità giuridica, vale a dire: la

capacità di agire, di comunicare, e di partecipare al “traffico giuridico”<sup>3</sup>. Secondo questa lettura, l’attribuzione di personalità non richiederebbe coscienza o sentimento, ma solo capacità funzionali analoghe a quelle delle persone giuridiche.

Nonostante i tanti, possibili, vantaggi, la proposta di introdurre un’*AI personhood* ha incontrato aspre reticenze e critiche sostanziali. Una prima obiezione riguarda la mancanza di interessi propri da tutelare: le persone giuridiche sono finzioni che servono a organizzare interessi umani collettivi e, ovviamente, l’IA non ha – né può avere – interessi autonomi. Attribuirle personalità significherebbe creare un soggetto fittizio che altri utilizzerebbero per i propri scopi, eludendo potenzialmente le eventuali responsabilità. Come ha osservato la *European Group on Ethics in Science and New Technologies*, conferire personalità giuridica ai *robot* potrebbe, per così dire, diluire la responsabilità umana e creare scappatoie legali<sup>4</sup>.

Una seconda obiezione concerne i diritti morali: se l’IA è autrice, dovrebbe avere il diritto alla paternità e all’integrità dell’opera. Ma come si eserciterebbero tali diritti? Un’IA può rivendicare di essere riconosciuta come autrice? Può opporsi a modifiche dell’opera? Chi agirebbe in suo nome? Domande che rivelano come l’estensione della personalità all’IA darebbe vita a tutta una serie di possibili paradossi pratici difficilmente risolvibili.

Una terza obiezione riguarda l’impatto sulla dignità umana. Riconoscere l’IA come persona significherebbe equipararla, almeno sul piano giuridico, agli esseri umani. Questo rischia di svilire il valore della personalità umana e di aprire la strada a forme di sfruttamento in cui le macchine vengono utilizzate per eludere obblighi che, invece, gravano sugli esseri umani. Chi scrive condivide tutte queste preoccupazioni e ritiene che l’attribuzione della personalità giuridica all’IA debba essere evitata, almeno se si tiene conto dello stato attuale della tecnologia e della riflessione etica.

Vien da sé che, poiché l’attribuzione di personalità giuridica all’IA risulta alquanto problematica, occorre individuare modelli alternativi che permettano di regolare la responsabilità per le attività dell’IA, riconoscendo e attribuendo dei diritti sulle opere generate dall’agente artificiale, senza necessariamente riconoscere all’IA una soggettività giuridica in senso proprio. Diverse proposte sono state avanzate in dottrina, ciascuna con vantaggi e svantaggi specifici.

<sup>3</sup> L.B. SOLUM, “Legal Personhood for Artificial Intelligences”, in *North Carolina Law Review*, 70(4), 1992, pp. 1231-1287.

<sup>4</sup> European Group on Ethics in Science and New Technologies, *Statement on Artificial Intelligence, Robotics and ‘Autonomous’ Systems*, Brussels, European Commission, 2018.

i) Una prima opzione consiste nell'applicare i principi tradizionali della responsabilità per fatto altrui o per le cose in custodia. In questa prospettiva, chi utilizza l'IA risponde dei danni eventualmente causati dal sistema, analogamente a come il proprietario di un animale risponde dei danni da esso provocati. Per quanto riguarda i diritti sulle opere generate artificialmente, essi spetterebbero all'utilizzatore in quanto la sua attività (fornire *input*, selezionare *output*) costituisce un contributo creativo rilevante. Questa soluzione ha il merito della semplicità e della continuità con il diritto vigente, ma solleva dubbi quando l'IA agisce in modo imprevedibile o quando l'intervento umano è minimo.

ii) Una seconda opzione prevede la creazione di fondi di compensazione finanziati dalle imprese che sviluppano e commercializzano sistemi di IA. Più in particolare, i danni causati dall'IA verrebbero risarciti attingendo a tali fondi, senza necessità di provare la colpa di un soggetto specifico. Questa soluzione, ispirata ai sistemi di responsabilità oggettiva e ai fondi di garanzia esistenti in altri settori (come quello dell'energia nucleare, oppure dei danni ambientali), risolve il problema pratico dell'indennizzo delle vittime ma non affronta la questione dei diritti sulle opere generate.

iii) Una terza opzione, già menzionata nel Capitolo III, consiste nell'elaborare forme di tutela *sui generis* per le opere generate dall'IA, attribuendo la loro titolarità non a un autore-persona ma a chi ha sostenuto l'investimento nello sviluppo del sistema. Questa soluzione si ispira ai diritti *sui generis* per le banche dati previsti dalla Direttiva 96/9/CE: non si riconosce l'IA come autrice, ciononostante si tutela l'investimento economico e tecnologico attraverso la previsione di un diritto esclusivo di durata limitata.

In questa sede si propende per un approccio pluralista che combini elementi di queste diverse soluzioni a seconda dei contesti che, di volta in volta, vengono considerati. Per la responsabilità extracontrattuale, si ritiene preferibile un modello di responsabilità oggettiva o di fondi di compensazione che garantisca il ristoro delle vittime senza particolari irrigidimenti di sorta. Per i diritti sulle opere generate dall'IA, si propone un regime differenziato che distingua tra diverse situazioni: intanto, quella delle opere create con l'ausilio dell'IA (protezione piena, titolarità all'utilizzatore); poi di quelle opere che sono co-create da esseri umani e IA (protezione ridotta, co-titolarità); ed infine delle opere generate autonomamente (tutela sui generis o dominio pubblico).

## 2. *L'attribuzione dei diritti nell'ecosistema dell'IA generativa*

La questione della soggettività giuridica dell'IA, esaminata nel paragrafo precedente, si intreccia strettamente con quella dell'attribuzione dei diritti sulle opere generate mediante il suo impiego. Se, infatti, l'IA non può essere considerata alla stregua di un'autrice, occorre individuare chi sia il titolare dei diritti nascenti dalla realizzazione dell'opera, vale a dire: il programmatore che ha sviluppato il sistema, l'utilizzatore che ha fornito l'*input*, l'impresa proprietaria del modello, oppure nessuno di essi? E ancora, nel caso di processi creativi complessi che coinvolgono molteplici attori umani e artificiali, come dovrebbero essere ripartiti i diritti?

Di qui la necessità di esaminare queste questioni, analizzando le diverse opzioni normative e i loro rispettivi effetti sul sistema della proprietà intellettuale.

Come si vedrà, nessuna delle soluzioni proposte è scevra da problemi. Nel dettaglio: l'attribuzione all'utilizzatore richiede di specificare quale grado di intervento sia sufficiente per qualificare qualcuno come autore; l'attribuzione al programmatore o all'impresa proprietaria rischia di concentrare eccessivamente i diritti in capo a pochi soggetti economicamente potenti; l'assenza di tutela (dominio pubblico) elimina gli incentivi all'innovazione. L'idea è quella di elaborare criteri flessibili che permettano di bilanciare le diverse esigenze in gioco, mediante un approccio adattivo e ponderato.

La domanda "chi è l'autore?" presuppone che ci sia un autore identificabile. Ma questa presupposizione potrebbe tranquillamente essere falsa: in alcuni casi, specialmente quando l'IA genera *output* autonomamente senza alcun intervento umano significativo, potrebbe non esservi un autore nel senso tradizionale del termine. Il diritto deve allora decidere se forzare la realtà riconducendola entro il perimetro (talvolta angusto) delle categorie esistenti (così da individuare sempre e comunque un autore) oppure se ammettere che possono anche esistere opere senza autore che necessitano di essere regolate di conseguenza.

Analogamente, nel caso *Thaler v. Perlmutter*<sup>5</sup>, l'High Court of Justice ha confermato che il diritto d'autore nel sistema di *common law* richiede un autore umano, negando protezione a un'immagine generata dal sistema *Creativity Machine* senza diretto intervento creativo dell'umano. Nel contesto britannico, di contro, permane la disciplina delle *computer-generated works*, in base alla quale, ai sensi della section 9(3) del *Copyright, Designs*

<sup>5</sup> *Thaler v. Secretary of State for Digital, Culture, Media and Sport*, High Court of Justice (England and Wales), EWHC 2412 (Admin), decisione del 21 settembre 2023.

and Patents Act 1988<sup>6</sup>, autore dell'opera è considerata la persona che ha posto in essere gli *arrangements necessary for the creation of the work*.

Se si ammette che deve sempre esservi un autore umano, occorre determinare chi sia tale autore tra i diversi soggetti coinvolti nel processo generativo. Le opzioni principali sono essenzialmente tre: il programmatore del sistema; l'utilizzatore che fornisce l'*input*; oppure (in alcuni sistemi giuridici) il proprietario del sistema come titolare derivato.

L'attribuzione al programmatore si giustifica considerando che egli ha creato lo strumento che rende possibile la generazione. Tuttavia, questa soluzione è problematica perché renderebbe il programmatore titolare di tutte le innumerevoli opere generate dal suo sistema, anche quando esse riflettono le successive scelte degli utilizzatori. L'attribuzione all'utilizzatore, invece, riconosce che il contributo creativo rilevante è l'ideazione e la selezione, non l'esecuzione tecnica. Ma occorre specificare quale grado di intervento umano sia sufficiente.

La proposta è quella di adottare il criterio del controllo creativo significativo in base al quale l'utilizzatore acquisisce la titolarità dei diritti nella misura in cui può dimostrare di aver esercitato un controllo creativo specifico sul processo generativo. Ad esempio, attraverso la specificità dei *prompt*, l'iterazione degli *output*, la selezione consapevole tra alternative, la modificazione del risultato. Al di sotto di tale soglia, la titolarità deve necessariamente essere attribuita al soggetto che ha investito nell'infrastruttura generativa (il proprietario del sistema), secondo una logica analoga al diritto *sui generis* sulle banche dati.

Una soluzione intermedia consiste nel riconoscere forme di co-autorialità tra i diversi soggetti che contribuiscono al processo generativo. Il diritto d'autore già conosce istituti di autorialità multipla: opere in collaborazione, opere composte, opere collettive. Tali categorie potrebbero essere adattate al contesto dell'IA generativa, distinguendo i contributi dei diversi attori.

Sorge però il dubbio in ordine a come gestire la co-autorialità con l'IA all'atto pratico. Questioni come l'esercizio dei diritti morali e la divisione dei compensi rimangono aperte. Soprattutto, ci si chiede se sia coerente parlare di collaborazione creativa quando uno dei soggetti coinvolti è un algoritmo privo di coscienza e volontà propria.

<sup>6</sup> *Copyright, Designs and Patents Act 1988*, c. 48, s. 9(3), testo consolidato ufficiale disponibile su [legislation.gov.uk](http://legislation.gov.uk). La disciplina britannica delle *computer-generated works* prevede che autore dell'opera sia la persona che ha posto in essere gli *arrangements necessary for the creation of the work*.

### 3. *Diritti fondamentali e IA*

L'analisi della soggettività giuridica dell'IA e dell'attribuzione dei diritti in ordine alle opere generate non può prescindere da una riflessione sui diritti fondamentali in gioco. Il diritto d'autore non opera in un vuoto normativo, ma si iscrive in un sistema costituzionale che tutela in prima battuta la dignità umana, la libertà di espressione, il diritto alla cultura e altri valori fondamentali. L'IA generativa solleva interrogativi circa il modo in cui questi diritti debbano essere interpretati e bilanciati nel nuovo contesto tecnologico.

Da qui, l'urgenza di esaminare tre dimensioni fondamentali dell'intersezione tra diritti fondamentali e IA. Innanzitutto, si analizzerà il ruolo della dignità umana come limite all'attribuzione di soggettività all'IA, sostenendo che riconoscere l'IA come persona rischia di svilire il valore dell'umano. In secondo luogo, si esaminerà la relazione tra libertà di espressione e creatività algoritmica, interrogandosi se l'IA debba godere di protezioni analoghe a quelle degli esseri umani. Infine, si affronterà la tensione tra diritto d'autore e diritto alla cultura nell'era dell'IA, chiedendosi come bilanciare l'accesso universale alla conoscenza con la tutela degli autori.

La dignità umana – riconosciuta come valore fondamentale dalle costituzioni moderne e dai trattati internazionali – costituisce il fondamento ultimo dei diritti umani. Essa esprime l'idea che ogni persona abbia un valore intrinseco e non strumentale<sup>7</sup>, e che meriti rispetto indipendentemente dalle sue capacità o da quelle che possono essere le sue prestazioni. L'articolo 1 della *Carta dei diritti fondamentali dell'Unione europea* afferma solennemente che “la dignità umana è inviolabile. Essa deve essere rispettata e tutelata”.

<sup>7</sup> Il concetto di valore intrinseco della persona – ossia di un valore che essa possiede in quanto tale, indipendentemente da qualunque uso o utilità che altri possano ricavarne – trova la sua formulazione filosofica più influente nella dottrina kantiana e nel principio secondo il quale si deve agire in modo da trattare l'altro sempre come fine e mai semplicemente come mezzo (I. KANT, *Fondazione della metafisica dei costumi*, trad. it., Milano, Bompiani, 2003). Questa formula stabilisce che la dignità umana non è commensurabile con alcun prezzo: ciò che ha un prezzo può essere sostituito con un equivalente; ciò che è al di sopra di ogni prezzo, e quindi non ammette equivalenti, possiede invece una dignità. Ne deriva che trattare un essere umano come mero strumento – ossia come entità valutabile esclusivamente in base alla sua utilità per scopi altrui – costituisce una violazione della sua dignità morale. Per una ricostruzione del concetto di persona nella tradizione bioetica e filosofico-giuridica italiana, si veda M. FARINA, voce “Persona”, in A.C. AMATO MANGIAMELI, G. SARACENI (a cura di), *Cento e una voce di Biogiuridica*, Torino, Giappichelli, 2024, pp. 347-351.

Riconoscere l'IA come soggetto di diritto – equiparandola giuridicamente alle persone fisiche – rischia di relativizzare la dignità umana. Se macchine prive di coscienza possono essere ritenute alla stregua di persone, allora la personalità non è più un attributo che distingue l'umano dal non-umano, ma diventa una qualificazione tecnica attribuibile a qualsiasi entità funzionalmente appropriata. Questo svilimento concettuale potrebbe avere conseguenze pratiche: facilitare lo sfruttamento degli esseri umani attraverso la mediazione tecnologica, erodere le protezioni giuridiche fondate sulla dignità, legittimare forme di trattamento che oggi consideriamo inaccettabili<sup>8</sup>.

Per queste ragioni, si ritiene che la dignità umana debba operare come limite invalicabile all'estensione della soggettività giuridica. L'IA può ricevere forme di riconoscimento normativo che ne regolino le attività e ne attribuiscono conseguenze giuridiche, ma non deve essere equiparata alle persone fisiche. Questa distinzione non è arbitraria, ma riflette una differenza ontologica che il diritto deve riconoscere e conservare: gli esseri umani sono soggetti morali dotati di dignità intrinseca, le macchine sono artefatti che possono acquisire rilevanza giuridica funzionale ma non equivalenza morale con le persone fisiche. Come precisato in precedenza, questa posizione non equivale a un antropocentrismo rigido che nega all'IA qualsiasi statuto normativo. Significa, piuttosto, che la dignità umana costituisce un tetto invalicabile, al di sotto del quale rimane lo spazio per forme graduate di riconoscimento giuridico calibrate sulle funzioni effettivamente svolte dai sistemi artificiali.

La libertà di espressione, garantita dall'articolo 11 della *Carta dei diritti fondamentali dell'Unione europea* e da analoghe disposizioni costituzionali nazionali, tutela, non solo il diritto di manifestare il proprio pensiero, ma anche il diritto di ricevere e comunicare informazioni. Nel contesto del diritto d'autore, la libertà di espressione opera come limite ai diritti esclusivi. Eccezioni come la citazione per critica, la parodia e l'uso a fini di informazione d'attualità sono giustificate dalla necessità di preservare il dibattito pubblico e la circolazione delle idee.

L'IA generativa solleva la questione se la libertà di espressione debba proteggere anche la creatività algoritmica. Se un sistema di IA produce contenuti che partecipano al discorso pubblico – articoli giornalistici, commenti politici, opere artistiche – tali contenuti meritano la protezione

<sup>8</sup> A.C. AMATO MANGIAMELI, *Natur@. Dimensioni della Biogiuridica*, Torino, Giappichelli, 2021. Il volume affronta il nesso tra natura umana, tecnologia e diritto, interrogandosi sui limiti etici e giuridici dell'ibridazione tra biologico e artificiale.

accordata alla parola umana? E se l'addestramento dei modelli su opere protette viene impedito per tutelare il diritto d'autore, si sta limitando indirettamente la capacità espressiva dei sistemi di IA?

La risposta a queste domande dipende da quale concezione della libertà di espressione si adotti. Se si ritiene che tale libertà tutela l'autonomia comunicativa degli esseri umani, allora l'IA non può invocarla direttamente, poiché manca di tale autonomia. Se invece si ritiene che la libertà di espressione protegga il valore sociale della circolazione delle informazioni indipendentemente dalla fonte, allora anche i contenuti generati da IA potrebbero meritare tutela. Si propende per la prima interpretazione che tuttavia non conduce all'esclusione dei contenuti generati da IA da qualsiasi forma di protezione costituzionale. Come s'è detto, quando l'IA veicola l'intenzionalità mediata dell'utilizzatore umano – attraverso un controllo creativo significativo sul processo generativo – la tutela della libertà di espressione si applica all'utilizzatore in quanto autore della comunicazione. I contenuti generati algoritmicamente con *input* umano rilevante beneficiano della protezione costituzionale attraverso la persona che li ha commissionati e diretti, non attraverso il sistema che li ha prodotti materialmente. Questa soluzione evita di dover risolvere la questione metafisica se i sistemi artificiali possano essere titolari di libertà fondamentali.

Un'attenzione particolare deve essere riservata al diritto alla cultura, sancito dall'articolo 27 della *Dichiarazione Universale dei Diritti Umani*, che oscilla tra la partecipazione attiva alla vita culturale e il godimento dei benefici del progresso scientifico. In modo particolare, nel contesto del diritto d'autore, tale principio funge da contrappeso ai monopoli privati, imponendo un bilanciamento costante tra la protezione degli autori e l'interesse pubblico all'accesso alla conoscenza.

L'avvento dell'IA generativa modifica profondamente i termini di questo equilibrio. Da un lato, la tecnologia può democratizzare la creazione culturale, abbattendo le barriere tecniche all'espressione; dall'altro, rischia di concentrare il potere culturale nelle mani delle poche imprese tecnologiche che controllano i modelli più avanzati. Per garantire che il diritto alla cultura non rimanga un'enunciazione astratta, è necessario guardare agli strumenti normativi già esistenti.

A livello europeo, la Direttiva (UE) 2019/790 (DSM) ha già tracciato un sentiero fondamentale introducendo eccezioni obbligatorie per il *text and data mining*. In particolare, l'articolo 3<sup>o</sup> prevede un'eccezione

<sup>9</sup> Direttiva (UE) 2019/790 del Parlamento europeo e del Consiglio del 17 aprile 2019 sul diritto d'autore e sui diritti connessi nel mercato unico digitale (c.d. Direttiva

specifica per finalità di ricerca scientifica, mentre l'articolo 4<sup>10</sup> stabilisce un'eccezione generale per qualsiasi finalità, lasciando tuttavia ai titolari dei diritti la possibilità di esercitare un *opt-out* per gli usi commerciali.

Anziché invocare un regime interamente nuovo, si propone di assumere questi strumenti come punto di partenza imprescindibile. L'approccio pluralista qui adottato mira a rafforzare tali eccezioni, integrandole con la promozione di modelli open source e con la proposta di destinare al dominio pubblico le opere generate autonomamente dall'IA. L'obiettivo è quello di trasformare l'innovazione tecnologica in un volano per la promozione e per la diffusione della cultura, evitando che si trasformi in una nuova forma di monopolio che limiti l'accesso universale al sapere<sup>11</sup>.

DSM), in *GUUE* L 130 del 17 maggio 2019, pp. 92-125. L'articolo 3 introduce un'eccezione obbligatoria al diritto d'autore per le attività di *text and data mining* effettuate da organismi di ricerca e istituti di tutela del patrimonio culturale, a condizione che le opere siano state lecitamente acquisite. L'eccezione è inderogabile: i titolari dei diritti non possono escluderla contrattualmente. La norma si applica esclusivamente a finalità di ricerca scientifica non commerciale, con esclusione degli usi commerciali anche quando condotti da soggetti di ricerca in partnership con imprese private.

<sup>10</sup> Direttiva (UE) 2019/790, cit., art. 4. La disposizione stabilisce un'eccezione generale per le attività di *text and data mining* condotte per qualsiasi finalità, incluse quelle commerciali. A differenza dell'art. 3, tuttavia, l'art. 4 è derogabile: i titolari dei diritti possono riservare espressamente i propri contenuti, esercitando il cosiddetto *opt-out* con modalità appropriate, anche in forma leggibile da macchina. La combinazione tra eccezione generale e *opt-out* è al centro del dibattito sull'addestramento dei modelli di IA generativa: laddove i titolari abbiano correttamente espresso la riserva, l'utilizzo delle loro opere per l'addestramento senza licenza integra una violazione del diritto d'autore, indipendentemente dall'uso trasformativo del risultato. Sul punto si veda G. PRIORA, "Il *text and data mining* nella Direttiva (UE) 2019/790: luci e ombre di una disciplina attesa", in *Diritto dell'informazione e dell'informatica*, 2020, pp. 419-452.

<sup>11</sup> Si tratta, in pratica, di quel presagio di un possibile monopolio di pochi, sul quale – fra i primi e in maniera incisiva – rifletteva già alcuni anni fa Bauman (Z. BAUMAN, *Modernità liquida*, trad. it., Roma-Bari, Laterza, 2002; *Vite di scarto*, trad. it., Roma-Bari, Laterza, 2004). Vien da sé che l'estensione di queste intuizioni all'ecosistema dell'IA generativa lascia presagire il rischio che si consolidi una frattura tra coloro i quali sono in grado di accedere agli strumenti di produzione e di fruizione della conoscenza algoritmica (i cosiddetti soggetti "di serie A") e coloro che, invece, rimangono esclusi da queste dinamiche e da questi vantaggi per ragioni economiche, culturali o tecnologiche (cioè, i soggetti "di serie B"). Inoltre, a proposito del nesso tra accesso alla conoscenza, disuguaglianza digitale e diritto d'autore nell'era dell'IA si vedano inoltre: G. PASCUZZI, *Il diritto dell'era digitale*, Bologna, il Mulino, 2020; T. PASQUALE, *The Coded Gaze: Bias in Artificial Intelligence*, Cambridge (MA), MIT Press, 2021.



PARTE TERZA

*Dimensioni etiche e scenari giuridici*



## CAPITOLO V

### ETICA DELL'IA E ORDINAMENTO GIURIDICO

SOMMARIO: 1. Fondamenti etici della regolamentazione dell'IA. – 2. La responsabilità etica nella creazione e nell'uso di sistemi IA. – 3. Questioni etiche specifiche nell'utilizzo creativo dell'IA. – 4. Verso un modello di ricerca etica e responsabile...

#### 1. *Fondamenti etici della regolamentazione dell'IA*

L'analisi condotta nei capitoli precedenti ha evidenziato come l'IA generativa sollevi tutta una serie di questioni giuridiche assai articolate e complesse, che non possono essere risolte semplicemente attraverso l'applicazione o l'adattamento delle norme esistenti. Tali questioni investono scelte fondamentali in ordine ai valori che l'ordinamento è chiamato a tutelare e ai principi che devono orientare la stessa regolamentazione dell'innovazione tecnologica. Prima di elaborare soluzioni normative specifiche, occorre dunque interrogarsi sui fondamenti etici che devono ispirare l'intervento del legislatore e l'azione adeguatrice dell'interprete nel confronto con l'IA. Quali principi morali devono innervare l'ossatura della *governance* dell'IA? Come si possono bilanciare valori potenzialmente confliggenti quali, ad esempio, l'innovazione e la precauzione, l'efficienza e la giustizia, la libertà e la responsabilità?

In questa sede – ed in questa precisa fase della trattazione – si guarderà alle dimensioni etiche e agli scenari giuridici, mostrando come la riflessione etica non rappresenti un'aggiunta opzionale bensì integri un presupposto necessario per ogni successivo intervento normativo sull'IA. In primo luogo, si analizzeranno i fondamenti etici della regolamentazione dell'IA, esaminando le principali teorie etiche (principialismo, etica della virtù, consequenzialismo, deontologia) nonché la loro applicabilità al contesto tecnologico. In secondo luogo, si affronterà la questione della responsabilità etica nella creazione e nell'utilizzo dei sistemi di IA, interrogandosi su come sia più opportuno distribuire le responsabilità nell'ambito di ecosistemi tecnologici complessi, che sono caratterizzati dalla presenza di molteplici attori. In terzo luogo, si esamineranno le questioni etiche connesse all'utilizzo creativo dell'IA, quali, a titolo d'esempio, il *training* su opere protette, i *deepfakes* e la mercificazione della creatività.

Come si vedrà, nessuna teoria etica isolata fornisce risposte univoche ai dilemmi posti dall'IA. Ciò che serve, piuttosto, è un approccio pluralista che sia in grado di integrare le diverse prospettive, riconoscendo che spesso principi e valori diversi possono entrare in tensione e richiedere bilanciamenti contestuali. L'etica non può – e non deve certo – sostituirsi al diritto, ma deve illuminare le scelte normative, evidenziandone le implicazioni e mettendo a disposizione criteri utili a valutarne la giustificabilità morale.

Il principlismo – elaborato da Tom Beauchamp e James Childress<sup>1</sup> nel contesto della bioetica medica – costituisce uno degli approcci più influenti nella regolamentazione etica delle tecnologie. Esso si basa su quattro principi fondamentali: *i*) quello dell'autonomia (rispetto della capacità degli individui di prendere decisioni informate); *ii*) quello della beneficenza (promozione del benessere); *iii*) quello della non maleficenza (evitare di causare danno); *iv*) quello della giustizia (equa distribuzione di benefici e oneri). Questi principi – seppur originariamente pensati per la medicina – sono stati estesi alla governance delle tecnologie emergenti, inclusa l'IA. Numerosi documenti internazionali – tra cui le *OECD AI Principles*<sup>2</sup> e la *UNESCO Recommendation on the Ethics of Artificial Intelligence*<sup>3</sup> – articolano valori analoghi (rispetto dell'autonomia, pro-

<sup>1</sup> T.L. BEAUCHAMP, J.F. CHILDRESS, *Principles of Biomedical Ethics*, 8a ed., New York, Oxford University Press, 2019.

<sup>2</sup> OECD, *Recommendation of the Council on Artificial Intelligence*, OECD/LEGAL/0449, adottata il 22 maggio 2019, aggiornata il 3 novembre 2023. Si tratta del primo atto intergovernativo in materia di IA, adottato dai 38 paesi membri dell'OCSE e, successivamente, da ulteriori otto Paesi non-membri (Argentina, Brasile, Egitto, Malta, Perù, Romania, Singapore e Ucraina), per un totale di 46 aderenti al dicembre 2023. I cinque principi fondamentali enunciati – crescita inclusiva e sviluppo sostenibile, rispetto dei valori umani e dell'equità, trasparenza e spiegabilità, robustezza e sicurezza, responsabilità – costituiscono il principale punto di riferimento per i quadri normativi nazionali e sovranazionali in materia di "IA affidabile". L'aggiornamento del 3 novembre 2023 ha riguardato principalmente la definizione di "AI system", riformulata per garantirne la pertinenza rispetto ai più recenti sviluppi tecnologici, inclusi i sistemi di IA generativa. Testo disponibile all'indirizzo: <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>.

<sup>3</sup> UNESCO, *Recommendation on the Ethics of Artificial Intelligence*, adottata per acclamazione da 193 Stati membri alla 41ª sessione della Conferenza Generale dell'UNESCO, Parigi, 23 novembre 2021, doc. SHS/BIO/PI/2021/1. Si tratta del primo strumento normativo internazionale dedicato all'etica dell'IA dotato di portata universale. La Raccomandazione individua quattro valori fondamentali – rispetto, protezione e promozione dei diritti umani e delle libertà fondamentali; prosperità dell'ambiente e degli ecosistemi; pace e convivenza; diversità culturale e inclusione – e undici principi operativi, tra cui trasparenza, responsabilità, *privacy*, equità e sostenibilità ambientale. Pur non essendo vincolante, il documento è indirizzato agli Stati membri con l'invito ad adottare quadri regolatori nazionali conformi ai principi enunciati e a riferire periodicamente sui progres-

mozione del benessere, prevenzione del danno, equità) pur senza derivarli esplicitamente dal *framework* di Beauchamp e Childress. In modo particolare, la struttura principialista offre uno strumento analitico per sistematizzare e mettere in relazione tali valori, non una dottrina ufficialmente adottata dagli organismi internazionali.

i) Il principio di autonomia, applicato all'IA generativa, richiede che gli individui mantengano il controllo sulle decisioni che li riguardano e non siano manipolati o ingannati dai sistemi algoritmici. Nel contesto del diritto d'autore, ciò implica che gli autori devono poter decidere liberamente se e come le proprie opere vengano utilizzate per l'addestramento dei modelli generativi, e che gli utilizzatori dei contenuti generati dall'IA devono essere informati circa l'origine artificiale degli stessi. La trasparenza diventa, così, un requisito derivato dal principio di autonomia. Del resto, solo conoscendo come funzionano i sistemi di IA e come vengono utilizzati i dati si può esercitare un controllo effettivo.

ii) Il principio di beneficenza impone di estendere e di massimizzare i benefici dell'IA per l'intera umanità. Nel campo della creatività, ciò significa favorire lo sviluppo di tecnologie che amplino le capacità espressive umane, democratizzino l'accesso agli strumenti creativi e arricchiscano il patrimonio culturale collettivo. La regolamentazione del diritto d'autore dovrebbe essere orientata, non solo a tutelare gli interessi economici degli autori, ma anche a promuovere l'innovazione artistica e la diffusione della conoscenza. Questo principio giustifica eccezioni ampie per usi trasformativi, sostegno allo sviluppo di modelli *open source* e politiche che facilitino l'accesso universale agli strumenti di IA creativa.

iii) Il principio di non maleficenza richiede di prevenire i danni potenziali dell'IA. Nel contesto creativo, tali danni includono: la violazione dei diritti degli autori attraverso appropriazione non autorizzata delle opere; la diffusione di contenuti falsi o manipolati (*deepfakes*) che ledono la reputazione o ingannano il pubblico; la perdita di posti di lavoro per artisti e creativi sostituiti da sistemi automatizzati; l'omologazione culturale derivante dalla predominanza di contenuti generati algoritmicamente secondo *pattern* standardizzati. La regolamentazione deve istituire salvaguardie appropriate contro questi rischi, attraverso norme sostanziali, meccanismi di controllo *ex ante* ed *ex post*, e sistemi di responsabilità e risarcimento.

iv) Il principio di giustizia, infine, impone un'equa distribuzione dei benefici e degli oneri dell'IA. Come si approfondirà nel prosieguo della

trattazione, l'economia dell'IA generativa è caratterizzata da asimmetrie profonde: poche imprese tecnologiche controllano i modelli più avanzati e catturano la maggior parte del valore economico generato, mentre i costi (sfruttamento dei dati degli autori, precarizzazione del lavoro creativo, impatti ambientali) sono distribuiti in modo trasversale e diffuso. Il principio di giustizia richiede politiche redistributive, atte a garantire che anche coloro che contribuiscono alla creazione di valore – autori i cui dati vengono utilizzati, lavoratori creativi – ricevano una parte equa dei benefici.

Il principlismo offre un *framework* utile ma non privo di limiti. Un primo problema è che i quattro principi possono entrare fra loro in conflitto: la tutela dell'autonomia degli autori (che potrebbero vietare l'uso delle proprie opere per l'addestramento) può ostacolare la beneficenza (sviluppo di tecnologie utili); la giustizia distributiva può richiedere limitazioni alla proprietà privata. Il principlismo non fornisce criteri gerarchici per risolvere tali conflitti, rimettendo piuttosto alla prudenza pratica il bilanciamento contestuale. Un secondo problema è che i principi, formulati in termini astratti, richiedono specificazione per essere operativi: che cosa significa concretamente "autonomia" nel contesto algoritmico? Quando un danno è evitabile e quindi viola la non maleficenza?

L'etica della virtù affonda le proprie radici nella filosofia aristotelica e che è stata ripresa nel dibattito contemporaneo da autori con prospettive divergenti, fra cui Alasdair MacIntyre, in chiave comunitarista e critica nei riguardi del liberalismo<sup>4</sup>, e Martha Nussbaum, in chiave liberale e universalista attraverso il *capabilities approach*<sup>5</sup>.

Anziché concentrarsi su principi astratti o regole universali, l'etica della virtù pone al centro la formazione del carattere morale degli agenti e lo sviluppo di disposizioni – le virtù – propedeutiche ad agire bene. Nel contesto tecnologico, ciò significa interrogarsi non solo su quali azioni siano giuste o sbagliate, ma anche su quali qualità caratteriali debbano possedere coloro che progettano, sviluppano e utilizzano i sistemi di IA.

<sup>4</sup> A. MACINTYRE, *After Virtue: A Study in Moral Theory*, 3rd ed., Notre Dame (IN), University of Notre Dame Press, 2007. L'Autore sviluppa una critica sistematica del progetto morale dell'Illuminismo e della frammentazione della moralità moderna, individuando nell'emotivismo una delle chiavi interpretative della cultura etica contemporanea e sostenendo il recupero della tradizione aristotelica delle virtù, radicata nelle pratiche, nelle tradizioni e nella dimensione comunitaria. La sua rilevanza per il contesto dell'IA generativa risiede nell'insistenza sulla formazione del carattere morale degli agenti come condizione di scelte eticamente fondate, in alternativa a una riduzione dell'etica a mera applicazione di regole o a semplice calcolo utilitaristico.

<sup>5</sup> M.C. NUSSBAUM, *Creating Capabilities: The Human Development Approach*, Cambridge (MA), Harvard University Press, 2011.

Da questo punto di vista, è bene notare sin da subito che alcune virtù per così dire classiche assumono un rilievo del tutto particolare nelle more dell'epoca algoritmica. Ad esempio, la prudenza (*phronesis*) – intesa come capacità di deliberare bene circa i mezzi appropriati per raggiungere fini buoni – è essenziale per orientarsi al meglio in situazioni di incertezza tecnologica nelle quali le conseguenze delle scelte non sono completamente prevedibili. I progettisti dei sistemi di IA, invero, devono esercitare la prudenza nel valutare i rischi, nel prevedere gli usi impropri, nel bilanciare innovazione e precauzione. La giustizia come virtù – distinta dalla giustizia come principio – consiste nella disposizione stabile a dare *a ciascuno ciò che gli è dovuto (unicuique suum tribuere)*<sup>6</sup>, evitando sia l'avidità che può sfociare anche nell'ingiusta appropriazione e sia l'eccessiva generosità che può distruggere gli incentivi.

La temperanza, da intendersi come autocontrollo e moderazione, si traduce nella capacità di resistere alla tentazione di sfruttare tecnologie potenti per fini moralmente dubbi anche quando ciò sarebbe economicamente vantaggioso. Nel contesto dell'IA generativa, ciò significa, ad esempio, astenersi dal generare *deepfakes* ingannevoli, guardarsi dal plagiare sistematicamente opere altrui attraverso l'addestramento non autorizzato, o dal diffondere contenuti tossici amplificati algoritmicamente. La fortezza morale – vale a dire, il coraggio – consiste nella capacità di difendere valori etici anche di fronte a pressioni economiche o competitive: emblematica, in tal senso, è la possibile scelta dell'impresa che rinuncia a pratiche moralmente dubbie, anche se concorrenti meno scrupolosi ne traggono vantaggio.

Particolarmente rilevante è la virtù dell'onestà intellettuale, che richiede di rappresentare accuratamente le capacità e i limiti dell'IA, senza cadere né nell'*hype* tecnologico, né nell'allarmismo ingiustificato. Molti problemi nell'interazione tra società e tecnologia derivano da aspettative irrealistiche generate da comunicazioni esagerate o fuorvianti. L'onestà intellettuale impone di comunicare chiaramente che cosa l'IA può e non può fare, quali rischi comporta e quali incertezze persistono.

L'etica della virtù offre anche strumenti per pensare la formazione etica dei professionisti tecnologici. Piuttosto che limitarsi a imporre regole esterne attraverso la legislazione, occorre promuovere lo sviluppo di una cultura professionale in cui virtù come prudenza, giustizia, tem-

<sup>6</sup> Sulla giustizia e, in modo particolare, circa il rapporto fra diritto e giustizia, cfr.: F. D'AGOSTINO, *Parole di giustizia*, Torino, Giappichelli, 2006; ID., *Giustizia. Elementi per una teoria*, Cinisello Balsamo (MI), San Paolo, 2006; ID., *Jus quia justum. Lezioni di filosofia del diritto e della religione*, Torino, Giappichelli, 2012.

peranza e onestà siano interiorizzate e costituiscano *standard* di eccellenza condivisi. Ciò richiede educazione etica nei *curricula* di informatica e ingegneria, codici deontologici professionali e meccanismi di *peer review* che valorizzino comportamenti virtuosi.

Tuttavia, l'etica della virtù presenta anch'essa limiti. Innanzitutto, essa presuppone agenti individuali dotati di carattere morale, ma molte decisioni nell'ecosistema dell'IA sono prese da organizzazioni complesse in cui le responsabilità sono frammentate. In secondo luogo, le virtù richiedono tempo per essere coltivate, mentre l'innovazione tecnologica procede a ritmi rapidissimi che non consentono la maturazione graduale di *habitus* etici consolidati. Infine, l'etica della virtù è meno adatta a fornire prescrizioni concrete su questioni specifiche di *policy*, limitandosi a indicare disposizioni caratteriali generali.

I due grandi capisaldi della filosofia morale moderna – ovvero, il consequenzialismo e la deontologia – offrono approcci radicalmente diversi alla valutazione etica delle azioni e delle politiche. Il consequenzialismo, nella sua versione più nota come utilitarismo, sostiene che la giustizia di un'azione dipende esclusivamente dalle sue conseguenze: è giusta l'azione che massimizza il benessere complessivo. La deontologia kantiana<sup>7</sup>, al contrario, sostiene che alcune azioni sono intrinsecamente giuste o sbagliate indipendentemente dalle conseguenze alle quali possono portare, perché rispettano oppure perché violano doveri fondamentali derivati dalla natura razionale dell'essere umano.

Applicato alla *governance* dell'IA generativa, il consequenzialismo suggerirebbe di valutare ogni scelta normativa in base ai suoi effetti sul benessere sociale complessivo. Ad esempio, il dilemma circa la necessità di proteggere o meno le opere generate dall'IA dovrebbe essere risolto chiedendosi quale regime normativo massimizza l'utilità sociale: la protezione incentiva l'investimento tecnologico incrementando la disponibilità di contenuti? Oppure limita l'accesso e riduce il benessere complessivo? L'analisi economica del diritto, profondamente influenzata dall'utilitarismo, adotta precisamente questo approccio, cercando di quantificare il bilancio costi/benefici delle diverse opzioni regolatorie.

Il consequenzialismo presenta vantaggi evidenti: fornisce un criterio apparentemente chiaro e oggettivo per le decisioni di *policy*, permette di considerare tutti gli effetti rilevanti e giustifica aggiustamenti normativi quando cambiano le circostanze empiriche. Tuttavia, esso si scontra

<sup>7</sup> I. KANT, *Fondazione della metafisica dei costumi*, trad. it., Milano, Bompiani, 2003.

con difficoltà teoriche e pratiche significative. Innanzitutto, l'idea stessa di misurare e di aggregare il benessere è estremamente complessa, specialmente quando sono in gioco valori incommensurabili come la libertà di espressione, la dignità degli autori, l'accesso alla cultura. In secondo luogo, il consequenzialismo può giustificare persino evidenti violazioni di diritti individuali, nel caso in cui ciò massimizzi l'utilità aggregata: ad esempio, potrebbe legittimare l'uso non autorizzato di opere protette se il beneficio sociale complessivo superasse il danno agli autori individuali.

La deontologia kantiana, dal canto suo, offre un approccio radicalmente diverso: anziché fondare la valutazione etica sulle conseguenze delle azioni, essa ricerca principi morali validi universalmente e indipendenti dagli effetti empirici. Nell'architettura dell'imperativo categorico, un'azione è moralmente lecita soltanto se la massima che la governa può essere universalizzata senza contraddizione. Vale la pena notare che anche la prima formulazione dell'imperativo – la formula della legge universale – conduce a conclusioni rilevanti per il diritto d'autore nell'era dell'IA: si potrebbe universalizzare la massima secondo cui “qualsiasi sistema di IA può essere addestrato su qualsiasi opera senza consenso dell'autore?” La risposta negativa è intuitiva: se tutti i produttori di contenuti adottassero quella massima, la produzione di opere originali cesserebbe di avere incentivi, autodistruggendo il presupposto stesso della prassi. L'argomento kantiano converge, così, con quello utilitarista sulla necessità di proteggere il consenso degli autori, ma per ragioni strutturalmente diverse. Applicato all'IA, questo principio vieta di trattare gli esseri umani come meri fornitori di dati o come strumenti per fini economici altrui. Gli autori non possono essere sfruttati per l'addestramento di modelli senza il loro consenso, anche se ciò generasse benefici sociali enormi, perché ciò violerebbe il loro *status* di fini in sé<sup>8</sup>.

<sup>8</sup> Evidente il rinvio all'imperativo categorico kantiano: “agisci in modo da trattare l'umanità, sia nella tua persona sia in quella di ogni altro, sempre anche come fine, mai semplicemente come mezzo” (I. KANT, *Fondazione della metafisica dei costumi*, cit.). Il principio esprime l'idea che ogni essere umano, in quanto dotato di razionalità e autonomia, possiede un valore assoluto che non può essere subordinato a nessun fine esterno: trattare una persona come mero mezzo – ossia sfruttarne le capacità o le produzioni senza riconoscerne la dignità autonoma – costituisce una violazione morale fondamentale. Applicato al contesto dell'IA generativa, tale principio vieta di utilizzare le opere degli autori come puri dati di addestramento senza il loro consenso informato: ciò li ridurrebbe a fornitori inconsapevoli di risorse per scopi economici altrui, in palese contraddizione con il rispetto della loro umanità. Per un approfondimento del principio kantiano nella sua rilevanza bioetica e giuridica si veda F. D'AGOSTINO, *Bioetica nella prospettiva della filosofia del diritto*, Torino, Giappichelli, 1996.

La deontologia offre fondamenti solidi per tutelare i diritti fondamentali e la dignità umana contro calcoli meramente utilitaristici. Tuttavia, anch'essa presenta dei limiti. Innanzitutto, i doveri assoluti possono entrare in conflitto fra di loro: il dovere di rispettare la proprietà intellettuale degli autori può collidere con il dovere di promuovere l'accesso universale alla cultura. Kant riteneva che tali conflitti fossero solo apparenti e risolvibili attraverso la ragione pratica, ma nella realtà normativa concreta i conflitti sono reali e richiedono scelte difficili. In secondo luogo, la deontologia tende a focalizzarsi su obblighi negativi (non violare diritti) piuttosto che su obblighi positivi (promuovere il benessere), risultando meno adatta a giustificare politiche redistributive o *welfare*.

Nella pratica della *governance* dell'IA, né il consequenzialismo puro, né la deontologia pura risultano pienamente adeguati allo scopo. Occorre, piuttosto, optare per un approccio integrato che consideri tanto le conseguenze delle politiche quanto il rispetto di principi e diritti fondamentali. In questa sede si adotta precisamente questo approccio di tipo pluralista, riconoscendo che la valutazione etica richiede l'integrazione di diverse prospettive teoriche piuttosto che l'adesione esclusiva a una tradizione filosofica. L'approccio pluralista si traduce in un metodo di riflessione equilibrata (*reflective equilibrium*<sup>9</sup>): si parte dai principi morali più solidi – il rispetto della dignità degli autori, il divieto di manipolazione, il diritto universale alla cultura – e si aggiustano reciprocamente e in corso d'opera, alla luce delle conseguenze concrete che si profilano all'orizzonte. In pratica, può dirsi che: i doveri deontologici fissano i vincoli, mentre il calcolo delle conseguenze guida le scelte tra opzioni normative che rispettano quei vincoli.

## 2. *La responsabilità etica nella creazione e nell'uso di sistemi IA*

La questione della responsabilità costituisce uno dei nodi più complessi della *governance* dell'intelligenza artificiale. Nei sistemi tecnologici

<sup>9</sup> J. RAWLS, *A Theory of Justice*, Cambridge (MA), Harvard University Press, 1971. Per la formulazione definitiva della teoria, si veda anche ID., *Political Liberalism*, New York, Columbia University Press, 1993. La giustizia come equità ("*justice as fairness*") si articola attorno a due principi fondamentali: il primo garantisce uguali libertà fondamentali per tutti; il secondo – il celebre principio di differenza – ammette disuguaglianze socioeconomiche solo nella misura in cui esse avvantaggino i membri meno favoriti della società. Applicato all'economia dell'IA generativa, tale principio impone di valutare se la concentrazione dei benefici dell'innovazione tecnologica nelle mani di pochi attori oligopolistici sia compatibile con i requisiti di equità distributiva.

tradizionali, la responsabilità per malfunzionamenti o danni è relativamente facile da attribuire: se un'automobile difettosa causa un incidente, ne risponde il costruttore; se un conducente la usa imprudentemente, ne risponde l'autista. Ma, nel caso di sistemi di IA generativa, specialmente quelli basati su *machine learning*, la catena causale è assai più complessa e coinvolge molteplici attori: ricercatori che sviluppano algoritmi, ingegneri che implementano il codice, imprese che addestrano modelli su *dataset* enormi, utilizzatori che forniscono *input* e, infine, i sistemi stessi che generano *output* in modi parzialmente inattesi e imprevedibili.

Di qui, la necessità di approfondire la questione della responsabilità etica nella creazione e nell'utilizzo di sistemi di IA da tre prospettive. Innanzitutto, bisogna volgere l'attenzione al concetto di responsabilità distribuita, evidenziando come – negli ecosistemi tecnologici complessi – la responsabilità non possa essere localizzata in un singolo attore, ma debba essere pensata come condivisa tra molteplici agenti. In secondo luogo, si affronterà il problema dell'*accountability*, ovvero il problema di chi deve rendere conto delle decisioni algoritmiche e secondo quali modalità. Infine, si esaminerà il ruolo della trasparenza e della spiegabilità intesi quali presupposti per l'attribuzione di responsabilità.

Il modello tradizionale di responsabilità individuale presuppone che vi sia un agente chiaramente identificabile le cui azioni causano un determinato evento dannoso. Questo modello funziona bene per situazioni semplici, ma diventa inadeguato nel momento in cui gli eventi sono il prodotto di interazioni complesse tra molteplici attori e fra vari sistemi. Come ha chiarito Helen Nissenbaum<sup>10</sup> nel suo lavoro pionieristico sull'etica computazionale, i sistemi socio-tecnici complessi rendono strutturalmente difficile attribuire la responsabilità a un singolo agente<sup>11</sup>. La causalità è distribuita tra progettisti, implementatori e utilizzatori, ignorare questa distribuzione ostacola, sia la giustizia verso le vittime, sia l'individuazione di soluzioni sistemiche efficaci.

Nel caso dell'IA generativa, contribuiscono al risultato finale: i ricercatori accademici che pubblicano architetture neurali innovative; le imprese che raccolgono *dataset* massivi e addestrano modelli; i program-

<sup>10</sup> H. NISSENBAUM, "Accountability in a Computerized Society", in *Science and Engineering Ethics*, 2(1), 1996, pp. 25-42.

<sup>11</sup> Sul nesso tra opacità dei sistemi algoritmici e difficoltà di imputare responsabilità giuridica per gli esiti dei processi decisionali automatizzati, v. G. GOMETZ, "Intelligenza artificiale, profilazione e nuove forme di discriminazione", *cit.*, il quale evidenzia come il funzionamento delle IA basate sul machine learning renda arduo ricostruire dall'esterno le ragioni del responso algoritmico e valutarne l'eventuale rilievo discriminatorio.

matori che scrivono il codice di inferenza; i *designer* che costruiscono interfacce utente; gli utilizzatori che forniscono *prompt*; e i sistemi stessi che, attraverso processi stocastici, generano *output* specifici.

Resta, ovviamente, da sciogliere il nodo della responsabilità giuridica derivante dagli *output* del sistema. Qualora un'opera generata violi il *copyright* di un autore terzo, è necessario stabilire criteri certi per l'individuazione del soggetto tenuto a rispondere del danno. Allo stesso modo, nel caso in cui il sistema produca contenuti diffamatori o discriminatori, occorre definire se la responsabilità debba ricadere sullo sviluppatore, sul fornitore del modello o sull'utente finale che ha istruito il *prompt*.

Il concetto di responsabilità distribuita riconosce che in tali contesti la responsabilità deve essere condivisa proporzionalmente al contributo causale e al controllo effettivo apportato/operato da ciascun attore. I ricercatori che sviluppano algoritmi hanno responsabilità per le capacità e i rischi intrinseci dei modelli proposti; le imprese che addestrano e commercializzano sistemi hanno responsabilità per la qualità dei dati, le salvaguardie implementate, e la comunicazione trasparente circa capacità e limiti; gli utilizzatori hanno responsabilità per l'uso appropriato degli strumenti. Questa distribuzione non significa che nessuno sia responsabile, ma che molteplici attori sono simultaneamente responsabili, ciascuno nella propria sfera di controllo.

Si badi, la responsabilità etica deve tradursi in *accountability* giuridica: vale a dire in meccanismi istituzionali che garantiscano che chi prende decisioni o chi compie azioni debba renderne conto e possa essere chiamato a rispondere. Nel contesto dell'IA, il richiamo al principio di *accountability* implica, innanzitutto, che si dia trasparenza con riguardo al funzionamento dei sistemi, alle decisioni progettuali assunte e, ovviamente, anche ai dati utilizzati. In secondo luogo, il rinvio alla nozione di *accountability* richiede la previsione di meccanismi di controllo – quali, ad esempio, *audit*, certificazioni, valutazioni d'impatto – che permettano di volta in volta di verificare la conformità a *standard* etici e giuridici. In terzo luogo, il riferimento all'*accountability* presuppone che si diano rimedi effettivi per chi subisce danni, come: l'accesso alla giustizia, la previsione di procedure di reclamo, così come di congrui risarcimenti.

Il Regolamento europeo sull'intelligenza artificiale<sup>12</sup> (*AI Act*) costituisce il primo quadro normativo orizzontale e vincolante sull'IA a livello mondiale, e rappresenta la risposta più compiuta finora elaborata

<sup>12</sup> Regolamento (UE) 2024/1689 del Parlamento europeo e del Consiglio del 13 giugno 2024 che stabilisce regole armonizzate sull'intelligenza artificiale

all'esigenza di istituire meccanismi strutturati di *accountability* nell'IA, attraverso obblighi di trasparenza, valutazione del rischio, conformità a standard tecnici e supervisione da parte di autorità nazionali e dell'*AI Office* europeo. La sua applicazione è progressiva: dal 2 febbraio 2025 si applicano le disposizioni generali, compresi i divieti per le pratiche a rischio inaccettabile e gli obblighi di alfabetizzazione all'IA; dal 2 agosto 2025 si applicano le regole di governance e gli obblighi per i modelli *GPAI* (*General Purpose AI*); dal 2 agosto 2026 entra in applicazione la maggior parte delle regole del Regolamento e prende avvio l'*enforcement*; dal 2 agosto 2027 si applicano anche le regole relative ai sistemi ad alto rischio integrati in prodotti regolamentati. Con più specifico riferimento al diritto d'autore e all'IA generativa, si propone l'introduzione di obblighi di *accountability* che includano: i) registri trasparenti dei *dataset* di addestramento; ii) meccanismi di *opt-out* per gli autori che non vogliono che le proprie opere siano utilizzate; iii) sistemi di *watermarking* e tracciabilità dei contenuti generati; iv) procedure di *notice-and-takedown* per contenuti illeciti; v) fondi di compensazione per risarcire autori danneggiati.

La trasparenza e la spiegabilità costituiscono presupposti fondamentali tanto della responsabilità etica quanto dell'*accountability* giuridica. Se, per un verso, è vero che non si può essere ritenuti responsabili per decisioni che non si comprendono; per un altro verso, è altrettanto vero che non si possono contestare decisioni algoritmiche di cui non si conoscono i criteri. Il principio di trasparenza richiede che sia accessibile informazione circa: quali dati vengono raccolti e utilizzati; come funzionano gli algoritmi; quali decisioni vengono automatizzate; chi controlla i sistemi. La spiegabilità, da par suo, va ben oltre la mera trasparenza: nel dettaglio, essa richiede che le decisioni algoritmiche possano essere spiegate in termini comprensibili agli esseri umani<sup>13</sup>.

Tuttavia, trasparenza e spiegabilità si scontrano con tutta una serie di difficoltà tecniche e commerciali. I modelli di *deep learning* più avanzati sono una sorta di scatole nere (*black boxes*) i cui meccanismi interni risultano opachi finanche ai loro stessi creatori. Richiedere spiegabilità completa potrebbe ostacolare l'innovazione. Inoltre, le imprese tecnologiche considerano gli algoritmi segreti commerciali e resistono a obblighi di *disclosure*. D'altro canto, non si può non tenere conto del fatto che il

<sup>13</sup> Cfr. G. SARTOR, *L'intelligenza artificiale e il diritto*, Torino, Giappichelli, 2022; A. SANTOSUOSSO, *Diritto, scienza, nuove tecnologie*, Padova, CEDAM, 2016; nonché M. FARINA (a cura di), *Lo "Scudo Giuridico Adattivo". Intelligenza artificiale, cybersecurity e diritti fondamentali nell'ecosistema serics*, Torino, Giappichelli, 2026.

bilanciamento tra trasparenza e protezione della proprietà intellettuale, tra spiegabilità e innovazione, costituisce una delle sfide centrali della governance dell'IA.

### 3. *Questioni etiche specifiche nell'utilizzo creativo dell'IA*

Oltre alle questioni etiche generali riferibili alla responsabilità e all'*accountability*, l'utilizzo creativo dell'IA solleva anche dei dilemmi specifici che meritano un'attenzione dedicata. Fra le aree più ostiche e particolarmente problematiche: l'addestramento di modelli su opere protette senza autorizzazione esplicita degli autori; la creazione di *deepfakes* e contenuti manipolati che mettono in discussione l'autenticità nell'era digitale; e la mercificazione della creatività attraverso la produzione algoritmica massiva di contenuti standardizzati. Ciascuna di queste questioni determina una tensione fra principi e valori fondamentali e, pertanto, richiede che si proceda, di volta in volta, ad un attento e ponderato bilanciamento tra le questioni chiamate in causa.

La questione etica e giuridica più controversa riguarda l'addestramento di modelli generativi su *dataset* che includono opere protette da *copyright* senza che vi sia stato un consenso esplicito da parte degli autori. Gli sviluppatori di IA sostengono che si tratta di un uso trasformativo e non consumativo, analogo a quello al quale si può procedere a seguito di una qualunque lettura umana. In pratica, così come un essere umano può leggere migliaia di romanzi e poi scrivere un proprio romanzo originale, allo stesso modo anche l'IA può apprendere da opere esistenti e successivamente generare creazioni nuove. Gli autori interessati, di converso, obiettano che si tratta di appropriazione massiva del loro lavoro creativo senza compenso o riconoscimento, violando tanto i diritti patrimoniali quanto quelli morali. Il parallelismo, tuttavia, non sembra poter essere accolto: a differenza del lettore umano, infatti, un modello di linguaggio può rievocare e ricombinare porzioni di testo in modo statisticamente estraibile e, in quest'ottica, alcuni studi di *memorization* hanno dimostrato che certi modelli riproducono *verbatim sequenze* delle opere di addestramento quando interrogati con *prompt* mirati<sup>14</sup> La questione empirica

<sup>14</sup> N. CARLINI, F. TRAMÈR, E. WALLACE, M. JAGIELSKI, A. HERBERT-VOSS, K. LEE, A. ROBERTS, T.B. BROWN, D. SONG, Ú. ERLINGSSON, A. OPREA, C. RAFFEL, "Extracting Training Data from Large Language Models", in *Proceedings of the 30th USENIX Security Symposium*, USENIX Association, 2021, pp. 2633-2650. Lo studio dimostra empiricamente che i modelli linguistici di grandi dimensioni ("*large language models*") possono memorizzare

di quanta informazione sia effettivamente “memorizzata” e riproducibile è rilevante per valutare se l’uso sia davvero non consumativo.

Dal punto di vista etico, la questione può essere formulata così: gli autori hanno un diritto morale a controllare l’uso delle proprie opere anche quando tale uso non comporta copia diretta? O il valore sociale dell’innovazione tecnologica giustifica limitazioni a tale controllo? Diverse teorie etiche conducono a risposte diverse. Un approccio consequenzialista potrebbe giustificare l’addestramento non autorizzato laddove i benefici sociali superassero i danni agli autori individuali. Un approccio deontologico kantiano potrebbe invece considerarlo una violazione del dovere di rispettare la proprietà altrui e la dignità degli autori come creatori.

Qui si propende per una posizione intermedia che riconosca tanto l’interesse legittimo degli autori quanto quello sociale all’innovazione. Nel dettaglio si propone di distinguere tra diversi tipi di addestramento: l’uso di opere per ricerca scientifica non commerciale dovrebbe beneficiare di ampie eccezioni; l’uso per sviluppo di modelli commerciali dovrebbe richiedere meccanismi di compensazione equa; l’uso che mira deliberatamente a replicare lo stile di autori specifici senza loro consenso dovrebbe essere vietato come appropriazione indebita. Questa gradazione riflette il riconoscimento che non tutti gli usi hanno la stessa rilevanza etica.

I *deepfakes* – ovverosia contenuti audiovisivi generati o manipolati attraverso l’IA in modo da risultare autentici pur essendo falsi – rappresentano una delle applicazioni più preoccupanti dell’IA generativa. Essi permettono di far dire o di far fare a persone cose che non hanno mai né detto né fatto, con livelli di realismo e di verosimiglianza tali da ingannare osservatori anche attenti. Gli usi potenziali dei *deepfakes* spaziano dall’intrattenimento innocuo (effetti speciali cinematografici) ad applicazioni gravemente dannose: pornografia non consensuale, diffamazione, manipolazione elettorale, frode finanziaria.

La questione etica centrale riguarda il valore dell’autenticità e della verità<sup>15</sup> in un’epoca in cui la tecnologia rende sempre più difficile distinguere il vero dal falso. Luciano Floridi, ha acutamente osservato che

e riprodurre verbatim sequenze di testo presenti nei dati di addestramento, inclusi contenuti potenzialmente sensibili o protetti da diritto d’autore, se interrogati con prompt opportunamente costruiti. Questo risultato ha implicazioni dirette per il dibattito sulla liceità dell’addestramento su opere protette: dimostra che l’apprendimento algoritmico non è puramente astrattivo come la lettura umana, ma comporta una forma di ritenzione testuale che può configurare riproduzione dell’opera originale.

<sup>15</sup> Sul valore epistemico e morale dell’autenticità e della verità nell’era digitale, con specifico riferimento alle tecnologie di manipolazione delle immagini e dei contenuti au-

quella attuale è l'età dell'Infosfera<sup>16</sup>, in cui la realtà è sempre mediata da rappresentazioni informazionali e la distinzione tra *online* e *offline* si assottiglia progressivamente fino a configurare un'esistenza *onlife*<sup>17</sup>. In questo contesto, i *deepfakes* non sono semplicemente inganni tecnologicamente sofisticati, ma rappresentano un attacco all'infrastruttura epistemica condivisa che rende possibile la comunicazione e la fiducia sociale<sup>18</sup>.

Da un punto di vista squisitamente giuridico, i *deepfakes* sollevano questioni che investono il diritto all'immagine, alla reputazione, alla *privacy*, oltre che di diritto d'autore quando vengono utilizzate opere protette come base per la manipolazione. Si sostiene la necessità di regolamentazione specifica che vieti la creazione e diffusione di *deepfakes* non consensuali dannosi, imponendo obblighi di *disclosure* quando contenuti manipolati vengono pubblicati e prevedendo sanzioni proporzionate alla gravità del danno. Al contempo, si riconosce che non tutti gli usi di tecnologie di sintesi sono eticamente problematici: la parodia, l'arte concet-

diovisivi, si vedano: C. LAZER *et al.*, "The Science of Fake News", in *Science*, 359(6380), 2018, pp. 1094-1096; H. FRANKFURT, *Sulla verità*, trad. it., Milano, Feltrinelli, 2007.

<sup>16</sup> Sul concetto di Infosfera si veda L. FLORIDI, *La quarta rivoluzione. Come l'infosfera sta trasformando il mondo*, Milano, Cortina, 2017. Con Infosfera Floridi designa l'ambiente complessivo – costituito da tutti gli agenti informazionali, le loro proprietà e le loro interazioni – nel quale gli esseri umani vivono sempre più immersi, fino al punto in cui la distinzione tra spazio *online* e *offline* si assottiglia progressivamente, configurando un'esistenza "onlife" in cui l'esperienza è strutturalmente mediata da rappresentazioni informazionali.

<sup>17</sup> Il termine "onlife" è coniato e sistematizzato in L. FLORIDI (a cura di), *The Onlife Manifesto. Being Human in a Hyperconnected Era*, Cham, Springer, 2015. Il concetto descrive la condizione in cui la distinzione tra vita *online* e *offline* si fa progressivamente irrilevante: gli esseri umani abitano un'unica realtà informazionale in cui l'esperienza è strutturalmente mediata da tecnologie digitali. In questo contesto, la manipolazione delle rappresentazioni informazionali – come nei *deepfakes* – non è un fenomeno marginale, ma un attacco alla struttura stessa dell'esperienza vissuta. Per l'elaborazione dell'Infosfera come ambiente in cui l'esistenza "onlife" si dispiega si veda L. FLORIDI, *La quarta rivoluzione*, cit., pp. 43-68.

<sup>18</sup> Il concetto di "agire comunicativo" è elaborato in J. HABERMAS, *Teoria dell'agire comunicativo*, trad. it., Bologna, il Mulino, 1986. L'agire comunicativo designa quella forma di azione sociale orientata all'intesa reciproca, in cui i partecipanti coordinano i propri piani d'azione non mediante coercizione o calcolo strategico, ma attraverso il raggiungimento consensuale di pretese di validità – di verità, di correttezza normativa, di sincerità. I *deepfakes* violano strutturalmente le condizioni di possibilità dell'agire comunicativo, falsificando la sincerità dell'enunciazione e manipolando le pretese di verità, essi non si limitano a ingannare singoli interlocutori, ma corrodono le basi fiduciarie su cui l'intera comunicazione pubblica si fonda. Sul rapporto tra etica della comunicazione habermasiana e nuove tecnologie si veda inoltre J. HABERMAS, *Un nuovo mutamento strutturale della sfera pubblica e la politica deliberativa*, trad. it., Milano, Raffaello Cortina, 2023.

tuale, gli effetti speciali meritano protezione nell'ambito della libertà di espressione.

Una critica ricorrente all'IA generativa riguarda il rischio di mercificazione della creatività: la riduzione dell'arte e della cultura a commodity prodotte industrialmente per massimizzare profitti, con conseguente perdita di valore spirituale, autenticità e significato profondo.

Questa preoccupazione ha radici profonde: Walter Benjamin<sup>19</sup>, analizzando la riproducibilità tecnica dell'arte, aveva colto l'ambivalenza del fenomeno. La perdita dell'aura – l'unicità irripetibile dell'opera originale – era per lui tanto una perdita culturale quanto un potenziale emancipatorio: la riproducibilità poteva sottrarre l'arte all'autorevolezza elitaria e avvicinarla alle masse. L'IA generativa radicalizza questa ambivalenza. Tuttavia, si potrebbe argomentare che la democratizzazione degli strumenti creativi costituisce un progresso, permettendo a persone prive di formazione artistica tradizionale di esprimersi. La storia mostra che ogni innovazione tecnologica ha suscitato resistenze da parte di *élite* culturali che vedevano minacciati i propri privilegi: la stampa fu accusata di svilire la cultura libraria manoscritta, la fotografia di distruggere la pittura, il cinema di corrompere il teatro. Col senno di poi, queste paure si sono rivelate infondate: le nuove tecnologie hanno ampliato le possibilità espressive piuttosto che distruggerle.

Si ritiene che occorra evitare tanto il luddismo culturale, che rifiuta *a priori* l'innovazione, quanto l'entusiasmo acritico, che ignora rischi reali. L'IA può essere uno strumento che amplifica la creatività umana o un meccanismo che la mercifica e la standardizza: quale esito prevarrà dipende dalle scelte normative, economiche e culturali che compiremo collettivamente. Garantire che l'IA serva l'umano fiorire piuttosto che

<sup>19</sup> W. BENJAMIN, *L'opera d'arte nell'epoca della sua riproducibilità tecnica*, trad. it., in ID., *Opere complete*, vol. VI, *Scritti 1934-1937*, Torino, Einaudi, 2004, pp. 3-45. Il concetto benjaminiano di "aura" designa l'insieme di qualità – unicità, autenticità, radicamento in un "*bic et nunc*" irripetibile – che caratterizza l'opera d'arte originale e che la riproduzione tecnica dissolve. Benjamin coglie nell'ambivalenza di questo fenomeno tanto una perdita – la scomparsa del valore culturale dell'opera, della sua "tradizione" – quanto una potenziale emancipazione: la riproducibilità tecnica sottrae l'arte all'aura elitaria e la consegna alle masse. L'IA generativa radicalizza questa tensione: essa non si limita a riprodurre opere esistenti, ma ne genera di nuove a costo marginale tendente a zero, azzerando ogni distinzione tra originale e copia e rendendo obsoleta la categoria stessa di "autorità dell'originale" su cui il diritto d'autore tradizionalmente si fonda. Cfr. anche A.C. AMATO MANGIAMELI, "Sul diritto d'autore e la proprietà intellettuale, ovvero l'opera dell'ingegno tra dono, diritto e tecnica", in EAD., *Informatica giuridica. Appunti e materiali ad uso di lezioni*, Torino, Giappichelli, 2015.

ridurre gli esseri umani a consumatori passivi di contenuti generati algoritmicamente richiede vigilanza etica continua.

#### 4. *Verso un modello di ricerca etica e responsabile...*

Dopo aver esaminato i fondamenti etici della regolamentazione dell'IA e le questioni specifiche sollevate dall'utilizzo creativo di tali tecnologie, è necessario soffermarsi più nel dettaglio sulla proposta che si intende avanzare nell'ottica di una ricerca etica responsabile nel campo dell'IA. L'idea di base è quella di non limitarsi a studiare l'IA dall'esterno, ma di incorporarla nella propria metodologia, assieme a principi e pratiche etiche, prospettando, così, un modello replicabile per altre iniziative di ricerca.

Uno degli obiettivi centrali è l'elaborazione di *guidelines* e *best practices* per la *governance* dell'IA generativa nel contesto del diritto d'autore. Tali *guidelines* si collocano in uno spazio intermedio tra norme giuridiche vincolanti e raccomandazioni puramente etiche, costituendo quella che è stata definita una sorta di *soft law*<sup>20</sup>, vale a dire: strumenti normativi privi di forza cogente ma dotati di autorevolezza morale e funzione orientativa. Le *guidelines* svolgono funzioni molteplici: *i*) forniscono interpretazioni autorevoli di norme esistenti; *ii*) anticipano sviluppi legislativi futuri; *iii*) offrono *standard* di condotta per operatori privati; *iv*) facilitano la convergenza internazionale in assenza di trattati vincolanti.

L'efficacia delle *guidelines* dipende da diversi fattori. In primo luogo, dalla loro qualità sostanziale: devono essere tecnicamente accurate, eticamente fondate, praticamente realizzabili. In secondo luogo, dalla legittimità procedurale del processo di elaborazione: devono coinvolgere *stakeholder* diversi, garantire trasparenza, incorporare meccanismi di revisione e aggiornamento. In terzo luogo, dai meccanismi di *implementa-*

<sup>20</sup> Sul concetto di *soft law* è bene ricordare la prolifica giurisprudenza sovranazionale secondo cui sebbene le raccomandazioni non siano vincolanti, i giudici nazionali sono tenuti a prenderle in considerazione per interpretare le norme interne o per chiarire il significato di altre disposizioni vincolanti dell'Unione; inoltre, nel parere consultivo sulla *Licetità della minaccia o dell'uso delle armi nucleari* (1996), la CIG ha riconosciuto che atti di *soft law* (come le risoluzioni dell'Assemblea Generale ONU) possono avere un valore normativo, servendo da prova per l'esistenza di una norma consuetudinaria in formazione o dell'emergere di un'*opinio iuris*; infine, la giurisprudenza europea (es. causa *Grecia c. Commissione*) ha chiarito che quando un'istituzione adotta linee guida e le rende pubbliche, essa autolimita il proprio potere discrezionale: l'inosservanza ingiustificata di tali *guidelines* può configurare una violazione del principio del legittimo affidamento e della parità di trattamento.

*tion*: adesione volontaria, certificazioni, incentivi reputazionali, eventuale incorporazione in contratti o regolamenti settoriali.

L'etica deve essere vista come un complemento necessario alla regolamentazione giuridica piuttosto che come suo sostituto. Il diritto, con la sua forza cogente e i suoi meccanismi di *enforcement*, rimane insostituibile per tutelare diritti fondamentali e prevenire abusi gravi. Tuttavia, il diritto è inevitabilmente in ritardo rispetto all'innovazione tecnologica, è relativamente rigido una volta promulgato, e non può catturare tutta la complessità delle situazioni concrete. L'etica<sup>21</sup>, attraverso la deliberazione riflessiva, la formazione del carattere e la promozione di culture organizzative virtuose, può colmare le lacune del diritto e favorire comportamenti responsabili anche in assenza di coercizione.

Questa complementarità si manifesta in diversi modi. Da un lato, l'etica può ispirare il diritto, fornendo i principi normativi che giustificano le scelte legislative. Dall'altro lato, il diritto può rafforzare il richiamo all'etica stessa, traducendo i principi morali in obblighi giuridici azionabili. Inoltre, talvolta, l'etica può supplire il diritto, guidando le condotte in aree non ancora regolamentate.

L'integrazione di competenze (filosofiche, giuridiche, tecniche, economiche ed etiche) è imprescindibile. Questa integrazione non si sviluppa spontaneamente ma richiede uno sforzo metodologico: creazione di linguaggi condivisi, individuazione di problemi comuni, costruzione di ponti concettuali tra discipline.

Da questo punto di vista, è opportuno mettere in atto diverse strategie: costituire *team* di ricerca multidisciplinari in cui giuristi e filosofi collaborino costantemente con informatici; organizzare *workshop* e

<sup>21</sup> Il termine "etica", nella sua accezione filosofica fondamentale, rimanda alla tradizione aristotelica che ne ha fissato i caratteri essenziali. In Aristotele, l'etica non è una disciplina meramente normativa che enuncia regole di condotta, ma una scienza pratica (*episteme praktike*) volta a chiarire in che cosa consista il bene per l'essere umano e come esso possa essere raggiunto attraverso la vita vissuta nella comunità politica. Il riferimento è in primo luogo ad ARISTOTELE, *Etica Nicomachea*, trad. it., Roma-Bari, Laterza, 2005 (in particolare il libro I, sulla felicità come fine dell'agire umano, e il libro II, sulla formazione delle virtù attraverso l'abitudine). Questa concezione pratica e comunitaria dell'etica si contrappone alle etiche normative moderne – consequenzialismo e deontologia – che astraggono il giudizio morale dalla vita concreta degli agenti. Il richiamo aristotelico serve qui a sottolineare che l'etica non si esaurisce nell'enunciazione di principi astratti, ma richiede la formazione di caratteri morali stabili (*ethos*) in grado di orientare la deliberazione pratica (*phronesis*) nelle situazioni concrete: condizione questa tanto più necessaria nell'epoca dell'IA, in cui la rapidità dell'innovazione tecnologica sfida continuamente la capacità delle norme di anticipare i problemi.

seminari che favoriscano il confronto tra prospettive diverse; realizzare pubblicazioni congiunte che dimostrino concretamente come l'interdisciplinarietà generi risultati impossibili da raggiungere attraverso approcci mono-disciplinari. Questo modello di ricerca può ispirare altre iniziative nel campo dell'etica dell'IA e oltre.

## CAPITOLO VI

### GIUSTIZIA DISTRIBUTIVA E IA GENERATIVA

SOMMARIO: 1. Teorie della giustizia applicate all'economia della conoscenza. – 2. Concentrazione del potere e asimmetrie nell'economia dell'IA. – 3. Diritti di remunerazione e modelli compensativi.

#### 1. *Teorie della giustizia applicate all'economia della conoscenza*

Nel precedente capitolo ci si è focalizzati sui fondamenti etici della regolamentazione dell'IA, concentrandosi sui principi normativi generali, sulle responsabilità e sulle questioni specifiche sollecitate al suo utilizzo creativo, ormai sempre più diffuso. A questo punto della nostra trattazione è utile porre l'accento su quella che è una dimensione etica particolarmente importante e critica, vale a dire, quella della giustizia distributiva nell'economia dell'IA generativa. Di qui, alcuni interrogativi tanto cruciali quanto ineludibili: come devono essere distribuiti i benefici e gli oneri dell'innovazione tecnologica affinché tale distribuzione possa essere considerata giusta? Chi deve ricevere compensi per il proprio contributo alla creazione di valore? Come garantire che l'accesso agli strumenti tecnologici non riproduca o amplifichi disuguaglianze preesistenti?

Tali domande non ammettono risposte puramente tecniche o economiche, ma richiedono il ricorso a teorie normative della giustizia. In questa sede, si esamineranno – seppur in maniera estremamente breve – tre tra le più influenti teorie della giustizia contemporanea, ossia: quella rawlsiana dell'equità, quella seniana delle capacità, e quella nozickiana dei diritti di proprietà. L'idea è quella di valutarne l'applicabilità all'economia della conoscenza e dell'IA. Invero, ciascuna teoria mette in evidenza aspetti diversi della questione, conducendo anche a prescrizioni normative parzialmente divergenti fra loro. L'intento è, dunque, quello di recuperare gli elementi nodali di ciascuna teoria, prospettando un approccio pluralista capace di integrare i contributi delle diverse prospettive.

In prima battuta, non si può non riprendere la teoria della giustizia come equità (*justice as fairness*)<sup>1</sup> elaborata da John Rawls è, senza alcun

<sup>1</sup> J. RAWLS, *A Theory of Justice*, Cambridge (MA), Harvard University Press, 1971. La teoria della giustizia come equità ("*justice as fairness*") si articola attorno a due principi

dubbio, uno dei contributi teorici più influenti per la filosofia politica del Novecento. Nel dettaglio, Rawls propone due principi di giustizia per regolare la struttura di base della società: il primo garantisce uguali libertà fondamentali per tutti; il secondo permette disuguaglianze economiche e sociali solo a condizione che queste massimizzino le aspettative dei membri meno avvantaggiati (principio di differenza) e siano associate a posizioni accessibili in condizioni di equa uguaglianza di opportunità.

Applicato all'economia dell'IA, il principio di differenza suggerisce che la protezione giuridica delle innovazioni e dei contenuti generati è giustificata solo se tale protezione agevola i membri più svantaggiati della società. Questo può avvenire se la protezione incentiva tecnologie accessibili a tutti, se i compensi agli innovatori finanziano servizi pubblici oppure se il sistema genera opportunità di impiego diffuse. Al contrario, laddove la protezione servisse esclusivamente ad arricchire alcune ristrette *élite* tecnologiche senza comportare benefici per la base sociale, essa violerebbe il secondo principio rawlsiano.

Si badi, nel *framework* rawlsiano, l'accesso alla conoscenza e alla cultura non figura esplicitamente tra i beni primari nella loro elencazione canonica (che include libertà, opportunità, reddito, ricchezza e le basi sociali del rispetto di sé). Tuttavia, esso risulta strettamente connesso all'equa uguaglianza di opportunità e alle basi sociali del rispetto di sé. La conoscenza, infatti, è la condizione necessaria per l'esercizio delle capacità morali e per la partecipazione alla vita democratica<sup>2</sup>. Ragion per cui, un sistema che concentri l'accesso alla conoscenza in mano di pochi,

scelti da individui razionali in una posizione originaria di ignoranza circa la propria collocazione sociale ("velo di ignoranza"): il primo garantisce uguali libertà fondamentali per tutti; il secondo – il principio di differenza – ammette disuguaglianze socio-economiche solo a condizione che esse massimizzino le aspettative dei membri meno avvantaggiati e siano associate a posizioni accessibili in condizioni di equa uguaglianza di opportunità. Per la riformulazione matura e la risposta alle critiche si veda anche ID., *Political Liberalism*, New York, Columbia University Press, 1993.

<sup>2</sup> L'affermazione che la conoscenza costituisca la condizione necessaria per l'esercizio delle capacità morali e per la partecipazione democratica trova un fondamento sistematico nella teoria dell'agire comunicativo di HABERMAS. Secondo HABERMAS, la razionalità comunicativa – che costituisce il presupposto della deliberazione democratica – richiede che i partecipanti al discorso siano in grado di comprendere, valutare e contestare le pretese di validità avanzate dagli altri: tale capacità dipende a sua volta dall'accesso a risorse cognitive e culturali adeguate. Cfr. J. HABERMAS, *Teoria dell'agire comunicativo*, trad. it., Bologna, il Mulino, 1986. Il medesimo nesso – tra accesso alla conoscenza, autonomia morale e partecipazione democratica – è ripreso e aggiornato alla luce delle trasformazioni digitali in ID., *Ein neuer Strukturwandel der Öffentlichkeit und die deliberative Politik*, Berlin, Suhrkamp, 2022, dove l'Autore osserva che la concentrazione del potere

pregiudicando le opportunità dei meno avvantaggiati, violerebbe il secondo principio di giustizia. E ciò, anche senza che la conoscenza venga formalmente classificata come un bene primario. Il bilanciamento tra l'incentivo all'innovazione e l'accesso universale deve quindi essere calibrato affinché la durata e l'ampiezza della protezione non creino barriere insormontabili, ma massimizzino i benefici complessivi per i membri più deboli della collettività.

Un'altra ricostruzione teorica che – nell'ambito di questa trattazione – si rivela quanto mai illuminante, è quella avanzata da Amartya Sen, che ha sviluppato l'approccio delle capacità (*capabilities approach*) come alternativa, tanto all'utilitarismo, quanto al rawlsianesimo contrattualista. Secondo Sen<sup>3</sup>, il metro della giustizia non è – e non dovrebbe essere – quello del benessere soggettivo né quello della distribuzione dei beni primari<sup>4</sup>, ma l'insieme delle capacità (*capabilities*) effettive che le persone hanno di vivere vite che hanno ragione di valorizzare<sup>5</sup>. Nell'alveo di que-

informazionale nelle mani di poche piattaforme private erode le basi cognitive della sfera pubblica, minacciando le condizioni strutturali della deliberazione democratica.

<sup>3</sup> A. SEN, *Development as Freedom*, New York, Alfred A. Knopf, 1999; ID., *Inequality Reexamined*, Oxford, Clarendon Press, 1992.

<sup>4</sup> Con l'espressione "beni primari" (*primary goods*) Rawls designa l'insieme dei beni che ogni persona razionale desidera, qualunque sia il proprio piano di vita: libertà fondamentali, opportunità, reddito, ricchezza e basi sociali del rispetto di sé. Nel quadro della sua teoria, tali beni costituiscono la metrica della giustizia distributiva, la misura attraverso la quale valutare se la situazione dei "meno favoriti" – i destinatari del principio di differenza – sia stata effettivamente migliorata. Cfr. J. RAWLS, *A Theory of Justice*, cit., pp. 54-65 e 78-81. La critica di SEN a questa metrica è che essa misura i mezzi anziché le libertà sostanziali: persone con uguale dotazione di beni primari possono trovarsi in condizioni radicalmente diverse di libertà reale, a seconda delle loro capacità individuali di convertire quei beni in funzionamenti effettivi. Una persona con una disabilità motoria e una sana, a parità di reddito e opportunità formali, non godono della stessa libertà sostanziale; la metrica rawlsiana dei beni primari non coglie questa asimmetria. La divergenza metodologica è più profonda: mentre RAWLS costruisce una teoria "trascendentale" – individua i principi di una società perfettamente giusta scelti da agenti razionali in condizioni ideali di ignoranza – SEN adotta un approccio comparativo e orientato alle realizzazioni concrete, con il quale si tratta non di definire l'ottimo ideale ma di valutare quali tra le alternative realmente disponibili riduca maggiormente le ingiustizie esistenti. Su questi punti si veda in particolare A. SEN, *The Idea of Justice*, Cambridge (MA), Harvard University Press, 2009; ID., *Inequality Reexamined*, cit.

<sup>5</sup> Il concetto di *capabilities* è il nucleo dell'approccio elaborato in A. SEN, *Development as Freedom*, cit. Le *capabilities* (capacitazioni) indicano l'insieme delle opportunità reali, delle libertà sostanziali e delle risorse di cui una persona dispone per condurre la vita che desidera e a cui attribuisce valore. Non ciò che la persona ha, né ciò che si sente, ma ciò che essa è realmente in grado di fare e di essere. Le *capabilities* non vanno intese

sta lettura, una società giusta è quella che espande le capacità di tutti, rimuovendo quegli ostacoli che impediscono agli individui di realizzare i funzionamenti (*functionings*)<sup>6</sup> che considerano importanti.

Applicato all'IA generativa, il *capabilities approach* sposta il criterio di valutazione dalla mera disponibilità formale degli strumenti alla capacità effettiva di ciascuno di servirsene per esercitare creatività autonoma e produrre espressione culturale significativa. In modo particolare, l'IA può espandere tali capacità, permettendo a individui privi di formazione artistica tradizionale di produrre contenuti esteticamente sofisticati. Tuttavia, se l'accesso agli strumenti più avanzati è limitato da costi proibitivi, da competenze tecniche inaccessibili o, ancora, da barriere linguistiche e culturali, allora la promessa di democratizzazione rimane irrealizzata e si generano nuove forme di esclusione<sup>7</sup>.

Sen enfatizza, inoltre, l'importanza della libertà processuale oltre che dei risultati sostanziali. Non basta che le persone abbiano accesso ai contenuti generati da IA; occorre infatti che possano partecipare attivamente ai processi creativi, esercitando *agency* piuttosto che rimanendo consumatori passivi<sup>8</sup>. Politiche che promuovano *open source*, educazione

come abilità innate o competenze, bensì come possibilità effettiva di convertire risorse in *functionings* (funzionamenti), che rappresentano gli stati e le attività effettivamente realizzati (ciò che una persona fa o è, come essere sana, ben nutrita, istruita). In relazione ai *functionings*, le *capabilities* sono le opportunità (o libertà) di scelta e, quindi, l'insieme delle combinazioni di funzionamenti, tra le quali la persona può scegliere. Una società giusta, secondo Sen, non si limita a garantire risorse o utilità aggregate, ma espande concretamente le *capabilities* di ogni individuo, rimuovendo gli ostacoli – economici, culturali, istituzionali – che ne impediscono la realizzazione. Per l'elaborazione sistematica si veda anche Id., *The Idea of Justice*, cit.

<sup>6</sup> I *functionings* (funzionamenti) designano, nell'approccio seniano, gli stati dell'essere e i modi dell'agire che una persona riesce a realizzare concretamente: essere nutriti, essere sani, partecipare alla vita comunitaria, avere rispetto di sé. Si distinguono dalle *capabilities* in quanto i funzionamenti sono ciò che la persona effettivamente è e fa, mentre le capacità sono l'insieme delle opzioni reali, tra le quali essa può scegliere. Questa distinzione ha rilevanza normativa: una politica che garantisce un funzionamento senza garantire la libertà di sceglierlo non espande le capacità. Cfr. A. SEN, *Development as Freedom*, cit., cap. 3.

<sup>7</sup> Sul nesso tra IA generativa, accesso alla conoscenza e democrazia si vedano: K. CRAWFORD, *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*, New Haven, Yale University Press, 2021. La tesi centrale è che i sistemi di IA non siano tecnologie neutrali, ma incorporino scelte politiche ed economiche che distribuiscono potere in modo asimmetrico, favorendo i soggetti che li sviluppano e controllano a discapito di chi ne subisce gli effetti. Sul piano più specificamente democratico-deliberativo si veda anche C. SUNSTEIN, *#Republic: Divided Democracy in the Age of Social Media*, Princeton, Princeton University Press, 2017.

<sup>8</sup> Il concetto di libertà processuale – distinta dalla libertà sostanziale intesa come

digitale universale e alfabetizzazione critica rispetto alle tecnologie algoritmiche servono a espandere le capacità.

Infine, un'ulteriore visione che merita di essere richiamata è senza dubbio quella di Robert Nozick, che propone una teoria libertaria della giustizia, che contrasta radicalmente con quelle più redistributive di Rawls e di Sen<sup>9</sup>. Il nucleo dell'argomento nozickiano è contenuto in *Anarchy, State, and Utopia*: "la tassazione del reddito da lavoro è assimilabile al lavoro forzato"<sup>10</sup>, poiché sottrarre i risultati del lavoro altrui equivale a sottrarre ore di vita e a indirizzarne l'attività secondo fini non scelti dal soggetto. Va precisato, tuttavia, che Nozick si riferisce specificamente

mero accesso ai risultati – è sviluppato in A. SEN, *The Idea of Justice*, Cambridge (MA), The Belknap Press of Harvard University Press, 2009. L'Autore sostiene che una valutazione adeguata della giustizia non può limitarsi a misurare i risultati ottenuti, ma deve considerare i processi attraverso i quali tali risultati vengono raggiunti. La partecipazione attiva alle decisioni che riguardano la propria vita è essa stessa un componente irriducibile del benessere e dell'*agency* individuale.

<sup>9</sup> Il contrasto tra la teoria libertaria di Nozick e le teorie redistributive di Rawls e Sen è strutturale e investe sia i principi sia il metodo. Rawls e Sen condividono la premessa che la giustizia richieda una valutazione delle condizioni effettive degli individui e una qualche forma di correzione delle disuguaglianze ingiustificate (Rawls attraverso il principio di differenza, che ammette solo le disuguaglianze che massimizzano la posizione dei meno avvantaggiati; Sen attraverso l'espansione delle *capabilities* di ciascuno, con particolare attenzione alle asimmetrie nelle capacità di conversione dei beni in libertà sostanziale). Nozick rigetta entrambe le impostazioni sul piano dei principi fondamentali. Per lui la giustizia non riguarda la distribuzione degli esiti (*end-state principles*), ma esclusivamente la legittimità dei processi attraverso i quali la distribuzione si è formata (*historical principles*). Una distribuzione è giusta se, e solo se, è stata prodotta da acquisizioni e trasferimenti storicamente legittimi, indipendentemente da come appaia dal punto di vista dei risultati. Ne deriva che qualsiasi intervento redistributivo coercitivo – tassazione inclusa – viola i diritti individuali, anche se produce esiti più "equi" nel senso rawlsiano o amplia le *capabilities* nel senso seniano. Il punto di distanza più profondo da Sen è metodologico. Mentre Sen adotta un approccio comparativo che valuta le realizzazioni concrete e le ingiustizie rimediabili, Nozick costruisce una teoria di principi assoluti che non ammette deroghe in nome delle conseguenze. Cfr. R. NOZICK, *Anarchy, State, and Utopia*, New York, Basic Books, 1974, pp. 149-182; J. RAWLS, *A Theory of Justice*, cit., pp. 65-83; A. SEN, *The Idea of Justice*, cit., pp. 52-74. Per una discussione critica del confronto tra le tre teorie applicato alla proprietà intellettuale si veda J. HUGHES, "The Philosophy of Intellectual Property", in *Georgetown Law Journal*, 77(2), 1988, pp. 287-366.

<sup>10</sup> L'argomento è formulato in R. NOZICK, *Anarchy, State, and Utopia*, New York, Basic Books, 1974, p. 169, con le seguenti parole: "Taxation of earnings from labor is on a par with forced labor. Seizing the results of someone's labor is equivalent to seizing hours from him and directing him to carry on various activities". Nozick sostiene che lo Stato minimo (*minimal State*), limitato alla protezione contro la forza, il furto e la frode, è l'unico giustificabile e che qualsiasi intervento redistributivo coercitivo viola i diritti di proprietà degli individui.

al reddito da lavoro (*earnings from labor*) e non ai profitti d'impresa. L'estensione dell'argomento a questi ultimi – pur sostenuta da alcuni commentatori libertari<sup>11</sup> – è un'operazione controversa che non può essere attribuita direttamente a Nozick senza le dovute cautele. È inoltre opportuno ricordare che lo stesso Nozick, in *The Examined Life*<sup>12</sup>, si è in parte distanziato dalle posizioni più rigide della sua opera principale, riconoscendo che l'impianto teorico di *Anarchy, State, and Utopia* aveva forse trascurato le esigenze dell'interdipendenza umana e della solidarietà sociale<sup>13</sup>. Applicata all'IA, la teoria nozickiana suggerirebbe di proteggere rigorosamente i diritti di proprietà delle opere utilizzate per l'addestramento e degli algoritmi sviluppati dai programmatori, in quanto qualsiasi uso non autorizzato di tali risorse costituirebbe una violazione di diritti, indipendentemente dai benefici sociali che potrebbe generare. Siffatta applicazione della teoria nozickiana alla proprietà intellettuale non è da considerare priva di tensioni interne, autori di orientamento libertario – tra cui Boldrin e Levine<sup>14</sup> – hanno, infatti, argomentato che i monopoli conferiti dal diritto d'autore siano incompatibili con i principi

<sup>11</sup> La tesi nozickiana è stata sviluppata, in senso più radicale, da M.N. ROTHBARD, *For a New Liberty: The Libertarian Manifesto*, New York, Macmillan, 1973, il quale ha esteso la critica alla tassazione in termini sistematici; e da E. FESER, "Taxation, Forced Labor, and Theft", in *The Independent Review*, 5(2), 2000, pp. 219-235, che ne discute e ne difende le implicazioni libertarie. Nell'ambito del diritto e dell'economia si veda anche R.A. EPSTEIN, "One Step Beyond Nozick's Minimal State: The Role of Forced Exchanges in Political Theory", in *Social Philosophy and Policy*, 22(1), 2005, pp. 286-313. Nessuno di questi autori applica tuttavia esplicitamente l'argomento ai profitti delle imprese tecnologiche nell'economia dell'IA. Tale estensione è, pertanto, da intendere quale proiezione logica della premessa libertaria e non una tesi direttamente attestata nella letteratura di riferimento.

<sup>12</sup> R. NOZICK, *The Examined Life: Philosophical Meditations*, New York, Simon & Schuster, 1989.

<sup>13</sup> Per una discussione critica dell'argomento si veda G.A. COHEN, *Self-Ownership, Freedom, and Equality*, Cambridge, Cambridge University Press, 1995.

<sup>14</sup> Cfr. M. BOLDRIN, D.K. LEVINE, *Against Intellectual Monopoly*, Cambridge, Cambridge University Press, 2008. La tesi centrale del volume è che brevetti e diritti d'autore costituiscano monopoli intellettuali che ostacolano la libera concorrenza anziché promuovere l'innovazione. La protezione giuridica esclusiva, lungi dall'incentivare la creazione, attribuisce ai titolari un potere di esclusiva sostenuto dall'ordinamento che limita la libertà degli altri di usare idee, tecniche e contenuti, comprimendo così la dinamica competitiva del mercato. La posizione degli Autori, di orientamento libertario e pro-mercato, è quindi interna al campo critico della proprietà intellettuale, non è la redistribuzione pubblica a essere contestata, ma il monopolio intellettuale privato garantito dallo Stato. Per una sintesi degli argomenti applicati specificamente ai brevetti, con evidenza empirica, si veda anche ID., "The Case Against Patents", in *Journal of Economic Perspectives*, 27(1), 2013, pp. 3-22.

della libera concorrenza, in quanto attribuiscono poteri coercitivi che limitano la libertà altrui di usare le proprie risorse.

Tuttavia, anche Nozick riconosce la clausola lockiana<sup>15</sup> della sufficienza, per cui l'appropriazione è legittima solo se lascia agli altri risorse sufficienti. In questa prospettiva, la proprietà intellettuale, come è già emerso nelle battute iniziali del presente lavoro, sembra porre seri problemi rispetto a tale clausola. Va inoltre considerato che Nozick ammette che, se l'acquisizione iniziale dei diritti fu ingiusta – ad esempio, nel caso di dataset costruiti mediante pratiche di raccolta (per es. attraverso *scraping*) dei dati non autorizzate, da intendere qui come esempio contemporaneo di acquisizione ingiusta —, allora la distribuzione attuale richiede rettificazione. Ne discende che anche nell'impianto nozickiano vi è spazio per interventi correttivi, pur entro confini teorici molto rigorosi.

## 2. Concentrazione del potere e asimmetrie nell'economia dell'IA

Le teorie della giustizia esaminate nel paragrafo precedente offrono criteri normativi per valutare la distribuzione dei benefici e degli oneri. Tuttavia, prima di applicare tali criteri occorre comprendere empiricamente in che modo il valore economico viene effettivamente distribuito nell'ecosistema dell'IA generativa. È, dunque, importante esaminare le dinamiche di concentrazione del potere economico e tecnologico, evidenziando asimmetrie profonde che sollevano preoccupazioni di giustizia distributiva.

Anzitutto, va detto che l'economia dell'IA è caratterizzata da tendenze oligopolistiche pronunciate: poche imprese – tra cui, al momento della reda-

<sup>15</sup> La clausola lockiana della sufficienza stabilisce che l'appropriazione privata di risorse è legittima solo se “lascia per gli altri abbastanza e altrettanto buono” (“*left enough, and as good for others*”): cfr. J. LOCKE, *Two Treatises of Government*, London, Awnsham Churchill, 1689. Nozick riprende e riformula la clausola in R. NOZICK, *Anarchy, State, and Utopia*, cit., pp. 174-182, sostenendo che l'appropriazione è legittima se non peggiora la situazione degli altri rispetto a una situazione di non-appropriazione. La clausola, trasposta ai giorni nostril, assume rilevanza cruciale nel contesto della proprietà intellettuale e dell'IA generative, per cui ogni nuova appropriazione di opere altrui per l'addestramento di modelli riduce le risorse informazionali liberamente accessibili e comprime lo spazio creativo disponibile per i futuri autori, sollevando interrogativi sulla compatibilità di tali pratiche con la condizione lockiana, tanto più che, a differenza delle risorse fisiche, le risorse informazionali non si esauriscono nel consumo ma si riducono nella loro accessibilità per effetto dell'appropriazione esclusiva. Per un'analisi critica della clausola lockiana applicata alla proprietà intellettuale si veda J. HUGHES, “The Philosophy of Intellectual Property”, in *Georgetown Law Journal*, 77(2), 1988, pp. 287-366.

zione, *OpenAI*, *Google*, *Meta*, *Microsoft* e *Anthropic* – controllano i modelli più avanzati e catturano la maggior parte del valore economico generato. L'elenco è necessariamente provvisorio in un settore in rapida evoluzione, ma la struttura oligopolistica sottostante è strutturale e non dipende dall'identità specifica degli attori dominanti in un dato momento. Questa concentrazione deriva da barriere all'entrata enormi: l'addestramento di modelli competitivi richiede investimenti in hardware computazionale (*GPU clusters*) dell'ordine di centinaia di milioni di dollari, accesso a *dataset* massivi e competenze tecniche rare. Le economie di scala e gli effetti di rete amplificano ulteriormente questa concentrazione: chi possiede più dati può addestrare modelli migliori, attirando più utenti e generando ancora più dati.

Di qui, un circolo che – oltre che vizioso – può diventare anche pericoloso se considerato nella prospettiva dell'accesso ai diritti e del loro effettivo esercizio. Ragion per cui, questa struttura oligopolistica solleva tutta una serie di questioni di giustizia e di democrazia. Se, difatti, poche imprese private controllano le tecnologie che mediano sempre più la produzione e il consumo culturale, è evidente che il loro potere diventa paragonabile a quello di *gatekeeper* tradizionali (editori, *broadcaster*) ma con capacità di influenza potenzialmente maggiori. Decisioni su quali contenuti amplificare, quali moderare, quali modelli di *business* adottare vengono prese da attori privati non democraticamente responsabili, con conseguenze per il pluralismo culturale e l'autonomia individuale<sup>16</sup>.

Nonostante la retorica della democratizzazione, l'accesso agli strumenti di IA creativa rimane assai limitato e stratificato. I modelli più avanzati sono accessibili solo a pagamento o attraverso *API* che richiedono competenze tecniche. Vi sono, poi, anche barriere linguistiche: la maggior parte dei sistemi funziona ottimamente in inglese, svantaggiando parlanti di altre lingue. Limitazioni culturali influenzano i risultati: i modelli addestrati prevalentemente su contenuti occidentali tendono a

<sup>16</sup> Il concetto di *nudging* – ossia di interventi sull'architettura delle scelte che orientano i comportamenti individuali verso esiti ritenuti preferibili senza ricorrere a proibizioni o incentivi economici – è sistematizzato in R.H. THALER, C.R. SUNSTEIN, *Nudge: Improving Decisions About Health, Wealth, and Happiness*, New Haven, Yale University Press, 2008. Applicato alle piattaforme di IA generativa, il *nudging* algoritmico solleva questioni di paternalismo digitale. In particolare, i sistemi di raccomandazione, i filtri e le modalità di presentazione dei contenuti orientano surrettiziamente le scelte degli utenti senza che questi ne siano consapevoli, comprimendo l'autonomia informata che le teorie democratiche presuppongono. Sul paternalismo digitale e i suoi limiti etici si veda anche C. SUNSTEIN, *The Ethics of Influence: Government in the Age of Behavioral Science*, Cambridge, Cambridge University Press, 2016; nonché G. BONGIOVANNI, G. SARTOR, C. VALENTINI (a cura di), *Reasonableness and Law*, Dordrecht, Springer, 2009.

riprodurre *bias*. Non da ultimo, persistono anche criticità economiche: chi può permettersi abbonamenti *premium* o *hardware* potente ottiene risultati migliori.

Questi divari digitali rischiano di amplificare le disuguaglianze sociali ed economiche preesistenti. Chi possiede risorse economiche, culturali e tecnologiche può sfruttare l'IA per potenziare ulteriormente le proprie capacità creative; diversamente, chi ne è privo rimane escluso, vedendo crescere il *gap* rispetto alle *élite* tecnologicamente avanzate. Garantire equità richiede politiche attive: accesso gratuito a strumenti di base, educazione digitale universale, sviluppo di modelli multilingue e culturalmente inclusivi, infrastrutture pubbliche di computazione.

Una strategia per contrastare la concentrazione oligopolistica consiste nel promuovere modelli *open source* che permettano a chiunque di accedere, studiare, modificare e redistribuire sistemi di IA. Iniziative come *Stable Diffusion*, *LLaMA* di Meta – distribuito con una licenza “*community*” che non rientra nella definizione OSI di *open source*, pur non coincidendo neppure, con il *software* proprietario, nel senso più tradizionalmente inteso – e progetti comunitari dimostrano la fattibilità tecnica di alternative al modello proprietario<sup>17</sup>. L'*open source* può favorire innovazione distribuita, trasparenza algoritmica, adattamento culturale, e riduzione di barriere all'accesso.

Tuttavia, l'*open source* solleva anche questioni complesse. Innanzitutto, chi sostiene i costi di sviluppo se i modelli sono gratuiti? Il finanziamento pubblico, le donazioni, o modelli di business basati su servizi accessori sono sufficienti? In secondo luogo, l'apertura del codice facili-

<sup>17</sup> *Stable Diffusion* è un modello generativo di immagini sviluppato da Stability AI e rilasciato nel 2022 con licenza *CreativeML Open RAIL-M*, che consente usi commerciali ma impone restrizioni sull'utilizzo per fini dannosi e richiede che le opere derivate mantengano la stessa licenza. Il codice sorgente è pubblicamente accessibile (“open weights”), ma ciò non equivale alla piena conformità alla definizione OSI (“*Open Source Initiative*”), che richiede anche libertà di modifica e redistribuzione senza restrizioni d'uso. *LLaMA* (*Large Language Model* Meta AI) è una famiglia di modelli linguistici sviluppata da Meta AI e distribuita con licenza “*community*” che consente usi commerciali con restrizioni per organizzazioni di grandi dimensioni: anche in questo caso non è corretto qualificarlo come *open source* in senso proprio secondo i criteri OSI, né come software proprietario in senso stretto. Questa zona grigia licenziale è rappresentativa di una più ampia tendenza del settore a sviluppare modelli ibridi – parzialmente aperti, parzialmente chiusi – che non rientrano ordinatamente nelle categorie tradizionali della proprietà intellettuale. Sul panorama delle licenze per modelli di IA si veda M. WIDDER, W. WEST, M. WHITTAKER, “Open (For Business): Big Tech, Concentrated Power, and the Political Economy of Open AI”, in *SSRN*, 2023.

ta anche usi malevoli: modelli *open source* possono essere impiegati per generare disinformazione, *deepfakes*, contenuti illegali. Come bilanciare trasparenza e sicurezza? Infine, l'*open source* non risolve automaticamente problemi di concentrazione: se mancano competenze e risorse computazionali, il mero accesso al codice non garantisce utilizzo effettivo.

### 3. *Diritti di remunerazione e modelli compensativi*

L'analisi delle asimmetrie distributive condotta nel paragrafo precedente evidenzia la necessità di ricorrere a meccanismi che garantiscano un'equa remunerazione agli autori le cui opere contribuiscono alla creazione di valore nell'ecosistema dell'IA. Il presente paragrafo esamina i fondamenti filosofici del diritto di remunerazione, analizza modelli compensativi esistenti o proposti, e valuta la loro adeguatezza ed efficacia.

Il diritto di remunerazione può essere "giustificato" ricorrendo a diverse teorie. Stando alla prospettiva lockiana, chi lavora creando opere merita compenso per il proprio sforzo, anche se le opere vengono successivamente utilizzate da altri in modi trasformativi. Adottando la prospettiva hegeliana, il riconoscimento economico dell'autore costituisce parte del riconoscimento della sua personalità manifestata nell'opera. Aderendo alla prospettiva utilitarista, la remunerazione incentiva la creazione futura. Optando per la prospettiva della giustizia distributiva, chi contribuisce alla generazione di valore economico deve partecipare equamente ai benefici.

Detto ciò, non v'è dubbio che determinare quale remunerazione sia effettivamente equa risulta assai complesso. L'utilizzo di un'opera per l'addestramento non sostituisce l'opera originale né ne riduce il valore di mercato diretto. Il contributo di ciascuna singola opera al valore del modello risultante è generalmente difficile da quantificare, e spesso modesto in termini assoluti. Tuttavia, la ricerca in *data attribution* e *machine unlearning* suggerisce che tale contributo non sia uniformemente trascurabile. Alcuni sottoinsiemi di dati di *training* possono, infatti, esercitare un'influenza sproporzionata sul comportamento del modello<sup>18</sup>. La difficoltà di misurazione non equivale, dunque, a irrilevanza causale, e ciò ha implicazioni significative per il calcolo di compensi equi. Come calcolare compensi appropriati in tali condizioni? Diverse proposte sono

<sup>18</sup> V. FELDMAN, *Does Learning Require Memorization? A Short Theory of a Long Tail*, New York, ACM Press, 2020; P.W. KOH, P. LIANG, *Understanding Black-box Predictions via Influence Functions*, La Jolla, International Conference on Machine Learning (ICML), 2017.

state avanzate: quella delle licenze collettive estese, quella dei sistemi di remunerazione proporzionale basati su utilizzo effettivo, quella delle tasse sulle imprese tecnologiche redistribuite agli autori.

Il *licensing* collettivo esteso (*Extended Collective Licensing - ECL*) è, invece, è un meccanismo già operativo in diversi Paesi nordici – in particolare in Danimarca (legge sul diritto d'autore), Norvegia e Svezia – per la gestione di diritti d'autore in contesti di uso massivo quali la radio-diffusione e la reprografia. Funziona così: un ente di gestione collettiva negozia licenze con utilizzatori per conto di tutti gli autori di una categoria, inclusi quelli non membri dell'ente; la licenza è vincolante per tutti, ma gli autori possono optare out; i compensi raccolti vengono distribuiti secondo criteri stabiliti. L'ECL potrebbe essere adattato all'IA: un ente negozierebbe licenze per l'addestramento, raccoglierebbe compensi dalle imprese, e li distribuirebbe agli autori le cui opere sono state utilizzate.

Questo modello presenta vantaggi: riduce i costi transattivi, garantisce i compensi anche ad autori che non potrebbero negoziare individualmente, permette l'utilizzo su larga scala necessario per l'addestramento. Tuttavia, solleva anche questioni: come determinare chi ha diritto a compensi quando i *dataset* includono miliardi di opere? Come calcolare la distribuzione equa? Come garantire trasparenza sull'utilizzo effettivo? Come gestire opere orfane o di autori non rintracciabili?

Da ultimo, una proposta più radicale, avanzata da alcuni teorici dell'economia della conoscenza, prevede l'introduzione di un reddito di base universale (*Universal Basic Income, UBI*) finanziato attraverso tassazione delle tecnologie che automatizzano il lavoro umano, incluso quello creativo. L'idea è che se l'IA sostituisce lavoro umano generando profitti per le imprese, parte di tali profitti dovrebbe essere redistribuita universalmente, garantendo a tutti i cittadini un reddito minimo indipendente dall'impiego.

Nel contesto creativo, l'UBI libererebbe artisti e autori dalla necessità di produrre per il mercato, permettendo espressione autentica. Tuttavia, l'UBI solleva obiezioni: è fiscalmente sostenibile? Non disincentiva il lavoro? Quali effetti macroeconomici avrebbe? E soprattutto, è la risposta appropriata ai problemi specifici del diritto d'autore, o serve piuttosto come soluzione generale a questioni più ampie di automazione e disoccupazione tecnologica? In realtà, pur riconoscendo il valore del dibattito sull'UBI, si ritiene che meccanismi più mirati e settoriali siano più immediatamente praticabili per affrontare le questioni di remunerazione nel contesto culturale.



PARTE QUARTA

*Prospettive regolative e traiettorie di ricerca*



## CAPITOLO VII

### MODELLI NORMATIVI PER L'IA GENERATIVA

SOMMARIO: 1. Approcci regolativi comparati. – 2. Principi filosofico-giuridici per una regolamentazione dell'IA creativa. – 3. Strumenti normativi innovativi.

#### 1. *Approcci regolativi comparati*

Sin qui, sono state analizzate le dimensioni filosofiche, giuridiche ed etiche dell'IA generativa con specifico riguardo al contesto del diritto d'autore. Si è reiteratamente posto l'accento sul fatto che l'IA pone sfide radicali ai fondamenti stessi del sistema autoriale, mettendo in discussione concetti nodali quali quello di autore, di originalità, di creatività e, pertanto, richiedendo un ripensamento profondo della distribuzione dei diritti e delle responsabilità nel quadro di una materia ben precisa, che, di per sé, è sempre stata fra le più complesse e articolate. A questo punto, è utile rivolgere lo sguardo alle soluzioni normative concrete, esaminando i modelli regolativi che sono stati già adottati – o che sono in fase di elaborazione – per tentare di governare l'uso dell'IA generativa in modo coerente con la tutela dei diritti, con le esigenze dell'innovazione e con le istanze di sicurezza pubblica.

Anzitutto si guarderà – in chiave comparativa – alle tre principali forme di regolamentazione, vale a dire: a) quella europea, oggi imperniata sull'*AI Act* e su una logica di graduazione del rischio, che nel caso italiano si intreccia altresì con un livello nazionale di attuazione e di specificazione; b) quella statunitense, maggiormente affidata a interventi settoriali, agli strumenti dell'Esecutivo e alla flessibilità interpretativa della *common law*, in specie del *fair use*; c) quella asiatica, che non può essere letta come un blocco omogeneo, ma che presenta modelli assai diversi tra loro, spaziando da forme di forte centralizzazione amministrativa a scelte più permissive o improntate alla *soft law*.

Si badi, l'analisi comparata – soprattutto in questa sede – non ha semplicemente un valore conoscitivo, ma permette di individuare tendenze e criticità comuni, nonché di cogliere dinamiche di convergenza e di divergenza regolativa. Nelle more di un'economia globalizzata in cui l'IA trascende confini nazionali, un'eccessiva frammentazione normativa può comportare costi elevati in termini di *compliance*, *forum shopping* e arbi-

traggio regolativo; d'altro canto, una convergenza solo apparente e troppo rapida rischierebbe di cristallizzare soluzioni inadeguate rispetto a una tecnologia ancora in evoluzione. Il problema, quindi, non è solo scegliere un modello tra quelli esistenti, ma comprendere quali valori ciascuno privilegi, quali strumenti utilizzi e quali costi sistemici sia disposto a sopportare.

Guardando all'Europa, va detto subito che l'Unione europea, nel 2024, ha adottato il Regolamento (UE) 2024/1689, meglio noto con l'acronimo di *AI Act*, vale a dire il primo quadro normativo orizzontale e compiuto dedicato all'intelligenza artificiale<sup>1</sup>. Il Regolamento è comunemente descritto come fondato su un approccio *risk-based*, poiché modula obblighi, divieti e cautele in ragione della diversa intensità del rischio che i sistemi di IA possono generare per la salute, la sicurezza e i diritti fondamentali. In questa prospettiva, il testo europeo combina pratiche vietate, discipline rafforzate per i sistemi ad alto rischio, obblighi di trasparenza per specifiche tipologie di sistemi e un'area residuale nella quale l'uso dell'IA non è soggetto a prescrizioni *ex ante* paragonabili a quelle previste per i casi più sensibili.

È tuttavia importante precisare che, per quanto riguarda l'IA generativa, il quadro europeo non può essere descritto correttamente limitandosi ad affermare che essa ricadrebbe senz'altro nel solo ambito del "rischio limitato". L'*AI Act*, infatti, da un lato prevede all'art. 50 obblighi di trasparenza per taluni sistemi e per taluni contenuti sintetici – si pensi, ad esempio, ai sistemi destinati a interagire direttamente con persone fisiche o ai contenuti artificialmente generati o manipolati, inclusi i c.d. *deepfakes* –; dall'altro lato, dedica il Chapter V ai *general-purpose AI models*, cioè ai modelli di uso generale, prevedendo per essi obblighi autonomi e, per i modelli con *systemic risk*, un ulteriore livello di disciplina<sup>2</sup>.

Per quanto concerne, in particolare, i *general-purpose AI models*, l'art. 53 impone ai *provider* un complesso di obblighi che risultano particolarmente rilevanti proprio ai fini del diritto d'autore. Essi comprendono: la predisposizione e l'aggiornamento della documentazione tecnica; la messa a disposizione di informazioni per i *provider* a valle che integrino il modello nei propri sistemi; l'adozione di una *policy* idonea ad assicurare il rispetto del diritto dell'Unione in materia di *copyright* e diritti connessi, anche con riguardo alla *reservation of rights* ai sensi dell'art. 4, par. 3, del-

<sup>1</sup> Regolamento (UE) 2024/1689 del Parlamento europeo e del Consiglio del 13 giugno 2024, recante regole armonizzate sull'intelligenza artificiale (*AI Act*).

<sup>2</sup> Regolamento (UE) 2024/1689, *cit.*, art. 50, Chapter V e art. 55; cfr. anche considerando 97, 101 e 111.

la Direttiva (UE) 2019/790; nonché la redazione e pubblicazione di una sintesi sufficientemente dettagliata del contenuto utilizzato per l'addestramento del modello<sup>3</sup>. Proprio questo punto mostra come il legislatore europeo abbia inteso affrontare il problema della trasparenza del *training* non già equiparando meccanicamente i sistemi generativi ai soli sistemi “a rischio limitato”, ma costruendo per essi un regime speciale, centrato sulla posizione del modello lungo la catena del valore.

Ora, va detto subito che l'approccio europeo riflette una tradizione in ossequio alla quale si propende per una regolazione di stampo precauzionale, che, dunque, privilegia la tutela dei diritti fondamentali e dei valori democratici anziché la sola massimizzazione dell'efficienza economica. In modo coerente con tale impianto, il Considerando 1 del Regolamento afferma che l'obiettivo è quello di promuovere un'intelligenza artificiale incentrata sull'essere umano e affidabile, garantendo al contempo un livello elevato di protezione della salute, della sicurezza e dei diritti fondamentali<sup>4</sup>. Tale orientamento – benché il termine “antropocentrico” non compaia nel testo finale nei termini in cui figurava nelle prime proposte – resta perfettamente coerente con la tradizione europea del diritto d'autore, da sempre attenta alla centralità della persona dell'autore, alla dimensione morale del diritto e alla necessità di bilanciare il progresso tecnico con la protezione dei creatori.

Non mancano, ovviamente, le critiche. Alcuni osservatori sostengono che gli obblighi previsti dal Regolamento siano eccessivamente onerosi e che essi possano, paradossalmente, favorire una concentrazione ulteriore del potere economico e tecnologico nelle mani dei grandi operatori, i quali dispongono di risorse per affrontare i costi di *compliance* assai più facilmente di quanto possano fare *start-up*, laboratori di ricerca e attori di dimensioni minori<sup>5</sup>. Altri, ancora, ritengono che la combinazione tra regole sui sistemi, regole sui modelli di uso generale e obblighi di trasparenza generi un sistema sofisticato ma non sempre semplice da ap-

<sup>3</sup> Regolamento (UE) 2024/1689, *cit.*, art. 53, par. 1, lett. a)-d); cfr. anche considerando 105 e 107, nonché Direttiva (UE) 2019/790, art. 4, par. 3.

<sup>4</sup> Regolamento (UE) 2024/1689, *cit.*, considerando 1.

<sup>5</sup> In questa prospettiva, v. R. KILIAN, L. JÄCK, D. EBEL, *European AI Standards – Technical Standardisation and Implementation Challenges under the EU AI Act*, in *European Journal of Risk Regulation*, 16, 2025, pp. 1038-1062, ove gli Autori, sulla base di interviste qualitative a ventitré organizzazioni operanti in diversi settori, mettono in evidenza come la combinazione tra tempistiche ristrette, costi di conformità e limitata partecipazione delle PMI ai processi di standardizzazione tecnica possa tradursi in barriere all'ingresso e in un vantaggio competitivo per gli operatori di maggiori dimensioni.

plicare<sup>6</sup>, soprattutto in presenza di modelli che vengono costantemente aggiornati, integrati e ridistribuiti attraverso rapporti contrattuali complessi<sup>7</sup>. Anche sul terreno dell'*enforcement* permangono questioni aperte: basti pensare alla concreta verificabilità delle informazioni rese dai *provider* circa il contenuto del *training* oppure all'effettività delle misure di gestione del rischio dichiarate<sup>8</sup>.

Per quanto concerne specificamente il diritto d'autore, poi, l'*AI Act* non risolve in via diretta le questioni fondamentali già menzionate nei capitoli precedenti. Lo stesso considerando 107 chiarisce che l'obbligo di pubblicare una sintesi del contenuto usato per il *training* ha finalità di trasparenza e di agevolazione dell'esercizio dei diritti da parte dei titolari, ma non equivale, di per sé, a una definizione legislativa definitiva della liceità dell'addestramento su opere protette. In termini più chiari, il Regolamento europeo costruisce un regime di *governance* e di informazione, ma lascia impregiudicata la questione dell'applicazione del diritto d'autore dell'Unione, inclusa la disciplina sul *text and data mining* di cui alla Direttiva (UE) 2019/790<sup>9</sup>. Ne deriva che il modello europeo, pur essendo più denso e strutturato di altri, non elimina il conflitto inter-

<sup>6</sup> Per una critica alla difficoltà applicativa di un sistema regolativo costruito su categorie di rischio rigide e su obblighi differenziati tra sistemi di IA, modelli di uso generale e regole di trasparenza, v. C. NOVELLI, F. CASOLARI, A. ROTOLO, M. TADDEO, L. FLORIDI, *Taking AI Risks Seriously: a New Assessment Model for the AI Act*, in *AI & Society*, 39, 2024, pp. 2493-2497.

<sup>7</sup> Commissione europea, *An agile Digital Rulebook for the EU*, in *Shaping Europe's digital future*, ove si osserva che la crescita del *digital acquis* ha esposto le imprese europee a *high compliance costs* e, in alcuni casi, a *legal uncertainty*; nello stesso documento, inoltre, si precisa che il futuro *Digital Fitness Check* sarà volto a studiare l'*interplay* tra le diverse regole digitali e il loro *cumulative impact on businesses*.

<sup>8</sup> Già con riguardo alla proposta di *AI Act*, v. M. BUSUIOC, D. CURTIN, M. ALMADA, *Reclaiming Transparency: Contesting the Logics of Secrecy within the AI Act*, in *European Law Open*, 2(1), 2023, pp. 79-105, ove si sottolinea come, in più punti, il soggetto regolato contribuisca in misura rilevante a determinare ciò che viene reso conoscibile all'esterno, con evidenti ricadute sulla trasparenza effettiva del sistema; nello stesso senso, v. D. HARTMANN, J.R.L. DE PEREIRA, C. STREITBÖRGER, B. BERENDT, *Addressing the Regulatory Gap: Moving Towards an EU AI Audit Ecosystem Beyond the AI Act by Including Civil Society*, in *AI and Ethics*, 5, 2025, pp. 3617-3638. Con specifico riguardo alla trasparenza sul *training*, v. inoltre M. KOWALA, *Protection of Press Publishers in the Age of Generative AI – In Search of Legal Remedies to Adapt to the Pace of Technology*, in *IIC - International Review of Intellectual Property and Competition Law*, 55, 2024, pp. 1604-1623.

<sup>9</sup> Direttiva (UE) 2019/790 del Parlamento europeo e del Consiglio del 17 aprile 2019 sul diritto d'autore e sui diritti connessi nel mercato unico digitale, spec. art. 4, par. 3; cfr. Regolamento (UE) 2024/1689, cit., considerando 107.

pretativo sul rapporto tra IA generativa e opere protette, ma lo inserisce entro una cornice di trasparenza e di responsabilizzazione più avanzata.

Se, tuttavia, ci si colloca dal punto di vista di uno Stato membro, appare evidente che il livello unionale non esaurisce da solo il quadro normativo. Il caso italiano lo dimostra con particolare chiarezza. Con la legge 23 settembre 2025, n. 132, il legislatore nazionale ha infatti costruito un secondo livello regolativo che, pur dovendo essere interpretato e applicato conformemente al Regolamento (UE) 2024/1689, interviene su profili organizzativi, di coordinamento istituzionale e di tutela sostanziale che l'*AI Act* lascia, in parte, alla competenza degli Stati membri. L'art. 1 della legge nazionale esplicita questa impostazione, promuovendo un utilizzo corretto, trasparente e responsabile dell'intelligenza artificiale in una dimensione antropocentrica; gli artt. 20 e 24, poi, disciplinano il versante istituzionale e quello dell'adeguamento dell'ordinamento interno, designando l'AgID e l'ACN quali Autorità nazionali per l'intelligenza artificiale, ferma restando, per i settori finanziari, la competenza di Banca d'Italia, CONSOB e IVASS, e delegando il Governo all'adozione di decreti legislativi di adeguamento all'*AI Act*<sup>10</sup>.

Per il tema specifico di questo lavoro, la disposizione nazionale di maggiore rilievo è però l'art. 25. Esso modifica il primo comma dell'art. 1 della legge 22 aprile 1941, n. 633, inserendo il riferimento all'opera dell'ingegno "umano" e precisando che la protezione può sussistere anche quando l'opera sia creata con l'ausilio di strumenti di intelligenza artificiale, purché costituisca risultato del lavoro intellettuale dell'autore. La scelta lessicale è significativa: il legislatore non equipara la macchina all'autore, ma chiarisce che l'uso dell'IA come strumento non esclude, di per sé, la tutela, purché il contributo creativo resti riconducibile all'essere umano. Lo stesso art. 25 introduce inoltre il nuovo art. 70-*septies* della legge sul diritto d'autore, il quale riconduce le riproduzioni e le estrazioni da opere o altri materiali, ai fini dell'estrazione di testo e dati mediante modelli e sistemi di IA, al regime già delineato dagli artt. 70-*ter* e 70-*quater*<sup>11</sup>.

Sul versante della tutela degli utenti e del contrasto alle manipolazioni,

<sup>10</sup> Legge 23 settembre 2025, n. 132, artt. 1, 20 e 24; cfr. anche Presidenza del Consiglio dei ministri – Conferenza Stato-città ed autonomie locali, *La normativa italiana: la legge 132/2025*, ultimo aggiornamento 29 dicembre 2025, ove si richiama la designazione dell'AgID e dell'ACN quali Autorità nazionali per l'intelligenza artificiale, ferma restando, per i settori finanziari, l'attribuzione alla Banca d'Italia, alla CONSOB e all'IVASS del ruolo di autorità di vigilanza del mercato ai sensi dell'art. 74, par. 6, dell'*AI Act*, nonché la delega al Governo per l'adozione, entro dodici mesi, di decreti legislativi di adeguamento.

<sup>11</sup> Legge 23 settembre 2025, n. 132, art. 25; cfr. legge 22 aprile 1941, n. 633,

il diritto italiano aggiunge ulteriori tasselli. L'art. 26 della legge n. 132 del 2025 interviene infatti sul codice penale e su altre disposizioni penali, introducendo, tra l'altro, il nuovo art. 612-*quater* c.p. in materia di illecita diffusione di contenuti generati o alterati con sistemi di intelligenza artificiale. Il dato sistemico è rilevante: mentre il livello europeo costruisce soprattutto obblighi di *risk management*, trasparenza e sorveglianza del mercato, il livello nazionale affianca a tali strumenti una risposta interna di tipo organizzativo, sanzionatorio e documentale. Anche per questa via si conferma che la responsabilità, nel governo dell'IA, non si misura più soltanto rispetto a norme astratte, ma sempre più rispetto alla capacità concreta di dimostrare, con documentazione verificabile, come siano stati gestiti il *design*, l'alimentazione dei modelli e il controllo dei possibili effetti degenerativi<sup>12</sup>.

È interessante mettere in evidenza che – rispetto all'Unione europea e, più in generale, rispetto al doppio livello europeo/nazionale che si osserva nel caso italiano – gli Stati Uniti hanno adottato un approccio molto diverso, che si connota per interventi regolatori meno sistematici e per l'affidamento prevalente a una combinazione di fonti legislative settoriali, poteri delle agenzie, atti dell'Esecutivo e *common law*. In altri termini, non esiste, allo stato, un quadro federale orizzontale paragonabile all'*AI Act*; la regolazione dell'IA negli Stati Uniti continua a svilupparsi attraverso strumenti tra loro eterogenei, in un contesto nel quale il Congresso ha discusso molteplici proposte senza però approdare, finora, a una disciplina generale unitaria<sup>13</sup>.

In tale contesto si colloca l'Executive Order 14110 del 30 ottobre 2023, intitolato *Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence*, con il quale l'amministrazione Biden ha cercato di imprimere un indirizzo coordinato e *Federal Government-wide* alla *governance* dell'IA. È però necessario essere precisi: non si trattava di una legge del Congresso né di un quadro generale direttamente applicabile ai privati in quanto tale. Il suo baricentro era rappresentato dalle direttive rivolte alle agenzie federali; nondimeno, alcune sue disposizioni, in particolare quelle della section 4.2, si fondavano su poteri legali pree-

art. 1, nel testo vigente dal 10 ottobre 2025, nonché art. 70-*septies*, introdotto dal medesimo art. 25.

<sup>12</sup> Legge 23 settembre 2025, n. 132, art. 26. Per un primo inquadramento delle modifiche penali introdotte dalla riforma, v. B. SCIALLA, *Le modifiche al Codice penale introdotte dalla normativa nazionale in materia di intelligenza artificiale* (L. 132/2025), in IRPA, 25 marzo 2026, spec. sul nuovo art. 612-*quater* c.p.

<sup>13</sup> Congressional Research Service, *Artificial Intelligence: Overview, Recent Advances, and Considerations for the 118th Congress*, R47644, 4 agosto 2023.

sistenti per imporre obblighi informativi a determinati sviluppatori di modelli avanzati e a soggetti dotati di rilevanti capacità computazionali. Tale scelta confermava l'elasticità ma, al contempo, anche la intrinseca precarietà degli strumenti esecutivi. Non a caso, l'Executive Order è stato poi abrogato il 20 gennaio 2025 tramite l'Executive Order 14148, del governo Trump, il che mostra quanto sia reversibile una forma di *governance* affidata in misura significativa al potere presidenziale in assenza di una disciplina congressuale organica<sup>14</sup>.

Nel quadro del diritto d'autore e dell'utilizzo dell'IA generativa, la dottrina del *fair use* costituisce il perno interpretativo centrale. Come è noto, il §107 del *Copyright Act* non prevede un catalogo chiuso di eccezioni, ma impone una valutazione *case by case* fondata su quattro fattori: lo scopo e il carattere dell'uso; la natura dell'opera protetta; la quantità e sostanzialità della porzione utilizzata; l'effetto dell'uso sul mercato dell'opera<sup>15</sup>. Proprio questa struttura aperta ha consentito al diritto statunitense di adattarsi storicamente ai mutamenti tecnologici con maggiore elasticità rispetto ai modelli più codificati; ma la stessa elasticità ha anche un costo, poiché rende inevitabilmente più incerta la previsione degli esiti applicativi.

Uno degli snodi concettuali più rilevanti, in questo contesto, è rappresentato dalla nozione di *transformative use*. A lungo essa è stata valorizzata quale chiave per riconoscere la liceità di usi che, pur implicando la riproduzione di opere preesistenti, risultassero giustificati da una funzione nuova, diversa o ulteriore rispetto a quella dell'opera originaria. È tuttavia necessario evitare semplificazioni. Nella decisione *Andy Warhol Foundation v. Goldsmith* del 2023, la Corte Suprema ha chiarito che la presenza di un nuovo "messaggio" o di una diversa percezione estetica non basta, di per sé, a risolvere il primo fattore in favore del *fair use*; la trasformatività deve essere valutata nel quadro più ampio della funzione economico-sociale dell'uso contestato e in rapporto agli altri fattori previsti dal §107<sup>16</sup>. Questo ridimensionamento è particolarmente importante nel dibattito sull'IA generativa, poiché impedisce di fare del solo carattere "nuovo" dell'*output* una scorciatoia argomentativa sufficiente a giustificare ogni forma di addestramento o di riuso.

<sup>14</sup> Executive Order 14110 of October 30, 2023, *Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence*, in *Federal Register*, 88 Fed. Reg. 75191 (1 novembre 2023); Executive Order 14148 of January 20, 2025, *Initial Rescissions of Harmful Executive Orders and Actions*, in *Federal Register*, 90 Fed. Reg. 8233 (28 gennaio 2025).

<sup>15</sup> 17 U.S.C. §107.

<sup>16</sup> *Andy Warhol Foundation for the Visual Arts, Inc. v. Goldsmith*, 598 U.S. 508 (2023).

Gli sviluppatori dell'IA generativa sostengono, non di rado, che l'addestramento su opere protette integri una forma di *fair use* trasformativo, poiché i modelli non si limiterebbero a memorizzare copie delle opere, ma estrarrebbero regolarità statistiche e rappresentazioni latenti funzionali a produrre *output* nuovi. Occorre, però, molta cautela. I precedenti spesso richiamati a sostegno di questa tesi – come *Kelly v. Arriba Soft* e *Perfect 10 v. Amazon*, relativi all'uso di *thumbnails* e di strumenti di ricerca per immagini, oppure *Authors Guild v. HathiTrust* e *Authors Guild v. Google*, concernenti la digitalizzazione di libri per finalità di ricerca testuale e accessibilità – offrono certamente analogie utili, ma non sono perfettamente sovrapponibili all'addestramento dei modelli generativi contemporanei<sup>17</sup>. In quei casi, infatti, le corti hanno valutato usi funzionalmente orientati alla ricerca, all'indicizzazione e alla consultazione di frammenti, non già sistemi capaci di generare nuovi contenuti sintetici in risposta a *input* dell'utente.

Di conseguenza, il modello statunitense presenta un vantaggio evidente in termini di flessibilità, ma sconta, specularmente, un alto grado di incertezza. Lo si vede chiaramente nel contenzioso ancora pendente sul rapporto tra *training* e *copyright*, fra cui spicca *Andersen et al. v. Stability AI Ltd. et al.*, ove si discute proprio se l'addestramento dei modelli generativi su grandi quantità di immagini protette senza licenza possa essere ricondotto a uso lecito oppure integri una forma di violazione massiva<sup>18</sup>. Finché questo contenzioso non maturerà in orientamenti consolidati, negli Stati Uniti l'equilibrio tra libertà di innovazione, tutela dei titolari dei diritti e trasparenza delle pratiche di addestramento resterà affidato, in larga misura, alla giurisprudenza e alla sua capacità di adattare categorie tradizionali a fenomeni tecnologicamente nuovi.

È necessario ricordare, inoltre, che l'approccio americano riflette una tradizione di diffidenza nei confronti degli interventi regolatori generali, percepiti spesso come ostacoli alla libertà economica e all'innovazione. La stessa *Copyright Clause* della Costituzione federale legittima la tutela autoriale in funzione della promozione del progresso della scienza e delle arti utili, e questa matrice strumentale continua a incidere profondamente sul modo in cui il sistema statunitense guarda al rapporto tra esclusiva, concorrenza e innovazione tecnologica<sup>19</sup>. Ne deriva una configurazione nella qua-

<sup>17</sup> *Kelly v. Arriba Soft Corp.*, 336 F.3d 811 (9th Cir. 2003); *Perfect 10, Inc. v. Amazon.com, Inc.*, 508 F.3d 1146 (9th Cir. 2007); *Authors Guild, Inc. v. HathiTrust*, 755 F.3d 87 (2d Cir. 2014); *Authors Guild v. Google, Inc.*, 804 F.3d 202 (2d Cir. 2015).

<sup>18</sup> *Andersen et al. v. Stability AI Ltd. et al.*, United States District Court for the Northern District of California, No. 3:23-cv-00201-WHO, filed January 13, 2023.

<sup>19</sup> U.S. Const., art. I, §8, cl. 8.

le il diritto d'autore non è tanto la sede di un riconoscimento personalistico dell'autore, quanto piuttosto uno strumento funzionale da bilanciare con gli interessi della diffusione culturale e dello sviluppo economico.

Guardando, poi, agli ordinamenti asiatici, va detto che questi presentano approcci diversificati, che riflettono contesti istituzionali, priorità politiche e tradizioni giuridiche profondamente differenti. Parlare genericamente di “modello asiatico” rischia, pertanto, di essere fuorviante. Più corretto è osservare che, all'interno dello spazio asiatico, convivono almeno tre direttrici: una prima, fortemente amministrativa e centralizzata, esemplificata dalla Cina; una seconda, più permissiva rispetto agli usi di dati e opere per finalità di analisi informativa, come avviene in Giappone; una terza, di carattere ibrido o sperimentale, che combina quadri legislativi generali, *linee guida* e strumenti di *governance* leggera, come si osserva in Corea del Sud, Singapore e India<sup>20</sup>.

L'approccio cinese è caratterizzato da una forte centralizzazione governativa e dalla priorità accordata alla stabilità sociale, alla sicurezza nazionale e al controllo del contenuto informativo. Le *Interim Measures for the Management of Generative Artificial Intelligence Services*, emanate nel luglio 2023 dalla *Cyberspace Administration of China* insieme ad altre autorità centrali, si applicano ai servizi generativi offerti al pubblico nel territorio della Cina continentale e definiscono come *provider* i soggetti che utilizzano tecnologie generative per offrire tali servizi, anche tramite interfacce programmabili. È però necessario essere precisi: la versione finale delle misure non introduce una licenza generale preventiva indistinta per tutti i *provider* di IA generativa. L'art. 17 rinvia infatti alla *security assessment* solo per i servizi dotati di *public opinion attributes* o *social mobilization capabilities*, secondo la normativa statale rilevante<sup>21</sup>. La Cina, dunque, non si limita a “vietare” o “consentire”, ma costruisce una griglia amministrativa che seleziona in modo mirato i casi ritenuti maggiormente sensibili per l'ordine pubblico informativo.

Il Giappone, da par suo, ha adottato un approccio significativamente più permissivo sul versante del diritto d'autore. L'art. 30-4 della legge giapponese sul *copyright* consente infatti l'utilizzazione di opere in tutti i casi in cui non sia scopo dell'utilizzatore godere, o far godere ad altri, i pensieri o sentimenti espressi nell'opera; tra tali casi la disposizione men-

<sup>20</sup> Per una ricognizione sintetica delle diverse traiettorie asiatiche, v. anche i materiali governativi richiamati nelle note seguenti.

<sup>21</sup> *Cyberspace Administration of China et al., Interim Measures for the Management of Generative Artificial Intelligence Services*, 10 luglio 2023, spec. artt. 1, 2 e 17.

zione espressamente la *data analysis*<sup>22</sup>. È pur vero che permane il limite dell'ingiustificato pregiudizio agli interessi del titolare, ma la struttura della norma rivela un'impostazione notevolmente più ampia di quella europea. In altri termini, il legislatore giapponese ha scelto di favorire l'analisi informativa e il trattamento computazionale dei contenuti, assumendo che in tali ipotesi l'interesse protetto dal *copyright* non coincida immediatamente con l'uso non espressivo dei materiali. Proprio per questo il Giappone è spesso richiamato, nel dibattito comparato, come esempio di ordinamento tendenzialmente favorevole all'addestramento di sistemi di IA mediante grandi masse di dati.

La Corea del Sud si colloca in una posizione intermedia. Da un lato, il legislatore ha approvato nel 2025 il *Framework Act on the Development of Artificial Intelligence and the Creation of a Foundation for Trust*, destinato a entrare in vigore nel 2026, con finalità esplicitamente connesse tanto allo sviluppo dell'IA quanto alla tutela dei diritti, degli interessi e della dignità dei cittadini; dall'altro lato, il sistema coreano si muove ancora entro una combinazione di promozione industriale, infrastrutture di fiducia e futura normazione secondaria. Nello stesso orizzonte di *governance* flessibile possono essere collocati, seppure con differenze proprie, anche Singapore e India: la prima ha privilegiato strumenti di *soft law* e di standardizzazione volontaria, come il proposto *Model AI Governance Framework for Generative AI* del 2024; la seconda ha sviluppato, soprattutto attraverso i documenti di NITI Aayog, un quadro di principi per la *Responsible AI* più che una disciplina legislativa unitaria<sup>23</sup>. Anche qui, dunque, la comparazione mostra che la partita regolativa si gioca non solo tra divieti e permessi, ma tra forme differenti di istituzionalizzazione del rischio e della fiducia.

Nel complesso, nessuno dei modelli considerati può essere ritenuto intrinsecamente superiore. L'Europa punta sulla densità normativa, sulla tutela dei diritti fondamentali e sulla responsabilizzazione degli attori della filiera, ma paga in termini di complessità applicativa e di costi di conformità. Gli Stati Uniti privilegiano la flessibilità, l'adattabilità giurisprudenziale e la libertà di innovazione, ma accettano, in cambio, un elevato livello di incertezza. Le giurisdizioni asiatiche, infine, offrono una

<sup>22</sup> *Copyright Act of Japan*, art. 30-4, nella traduzione ufficiale del *Japanese Law Translation Database System*.

<sup>23</sup> Republic of Korea, *Framework Act on the Development of Artificial Intelligence and the Creation of a Foundation for Trust*, Act No. 20676, 21 gennaio 2025, effective 22 gennaio 2026; IMDA - *AI Verify Foundation, Proposed Model AI Governance Framework for Generative AI*, 16 gennaio 2024; NITI Aayog, *Responsible AI for All: Strategy for India*, 2021.

pluralità di soluzioni che oscillano tra controllo amministrativo forte, eccezioni ampie per l'analisi dei dati e sperimentazioni di *soft law*. Da questo punto di vista, l'analisi comparata conferma che la *governance* dell'IA generativa non può essere ridotta a una scelta tra modelli precostituiti: essa richiede piuttosto un continuo lavoro di bilanciamento tra innovazione, trasparenza, sicurezza, tutela della persona e protezione della creatività, bilanciamento che ogni ordinamento struttura secondo la propria cultura costituzionale e la propria idea di giustizia tecnologica.

## 2. *Principi filosofico-giuridici per una regolamentazione dell'IA creativa*

L'analisi comparata degli approcci regolativi esistenti sin qui condotta mette in evidenza come – al di là delle peculiarità e delle differenze – esistano principi etico-giuridici trasversali, che possono orientare fruttuosamente la *governance* dell'IA generativa. Fra questi, spiccano senza dubbio: *a*) il principio di precauzione (che guida le scelte in condizioni di incertezza scientifica circa rischi potenzialmente gravi); *b*) i principi di sussidiarietà e proporzionalità (che delimitano l'ambito e l'intensità dell'intervento regolatorio); *c*) il principio di bilanciamento tra innovazione e tutela dei diritti esistenti. Si tratta di principi, che non forniscono risposte univoche, ma che, più che altro, individuano quei criteri normativi che disciplinano il processo deliberativo attraverso il quale le scelte regolatorie devono essere giustificate. Principi che, in questa sede, meritano di essere esaminati più nel dettaglio.

*Il principio di precauzione* – Codificato nel diritto ambientale europeo e internazionale, il principio di precauzione stabilisce che “*Where there are threats of serious or irreversible damage, lack of full scientific certainty shall not be used as a reason for postponing cost-effective measures to prevent environmental degradation*”<sup>24</sup>.

Applicato alla *governance* tecnologica, tale principio suggerisce che l'assenza di prove definitive di danno non giustifica l'inazione, anche quando esistono ragioni plausibili per ritenere che una tecnologia possa causare danni significativi.

Nel contesto dell'IA generativa, il principio di precauzione giustificerebbe misure preventive quali: obblighi di valutazione d'impatto *ex ante* prima del rilascio di sistemi potenti; meccanismi di supervisione e monitoraggio continuo; standard di sicurezza elevati anche in assenza di

<sup>24</sup> Dichiarazione di Rio sull'ambiente e lo sviluppo, Rio de Janeiro: ONU, 1992.

incidenti pregressi; inversione dell'onere della prova (chi sviluppa tecnologie deve dimostrare sicurezza, non chi teme danni deve provare pericolo). Queste misure riflettono l'idea che – date le potenzialità dell'IA – sia preferibile un eccesso di cautela piuttosto che permettere danni che poi risultano difficilmente rimediabili.

Tuttavia, il principio di precauzione è controverso. I critici sostengono che interpretato rigidamente esso condurrebbe a paralisi regolativa: poiché qualsiasi tecnologia comporta rischi ipotetici, si dovrebbe vietare o limitare qualsiasi innovazione. Inoltre, il principio ignora i costi dell'inazione: rinunciare all'IA significa rinunciare anche ai suoi benefici potenziali in medicina, educazione, ricerca scientifica. Una versione moderata del principio richiede che le misure precauzionali siano proporzionate alla gravità del danno temuto e alla plausibilità del rischio, e che siano sottoposte a revisione periodica alla luce di nuove evidenze.

Una via da seguire può essere quella di optare per un'applicazione contestuale e graduata del principio di precauzione. Per i sistemi di IA impiegati in contesti ad alto rischio (salute, sicurezza pubblica, diritti fondamentali), misure precauzionali stringenti risultano pienamente giustificate. Per le applicazioni creative a basso rischio, si dovrebbe invece propendere per un approccio permissivo in cui il *monitoring ex post* risulta più consono. Ovviamente, si tratta di una gradazione che richiede valutazioni caso per caso guidate da *expertise* tecnica e deliberazione democratica.

*b) Il principio di sussidiarietà e proporzionalità* – I principi di sussidiarietà e proporzionalità sono sanciti nell'art. 5 del *Trattato sull'Unione europea* (TUE). Nel dettaglio, il primo principio – ovvero quello di sussidiarietà – stabilisce che l'autorità superiore dovrebbe intervenire solo nel momento in cui gli obiettivi non possano essere realizzati adeguatamente a livello inferiore (nazionale o locale). Diversamente, il secondo principio – ossia quello di proporzionalità – richiede che le misure adottate non trascendano quelle strettamente necessarie per il raggiungimento degli obiettivi legittimi perseguiti.

Applicati alla regolamentazione dell'IA, questi stessi principi producono varie ricadute. Innanzitutto, suggeriscono di preferire sempre l'adozione di strumenti meno invasivi. Ragion per cui, prima di imporre obblighi vincolanti, occorre valutare se gli strumenti di *soft law*, di *self-regulation* del settore o i meccanismi di mercato possano essere sufficienti a garantire e a riequilibrare la situazione considerata. In secondo luogo, tali principi richiamano la necessità di procedere a una calibrazione degli obblighi. Ciò significa che, anziché imporre requisiti uniformi a tutti i si-

stemi, è opportuno differenziare gli interventi e le richieste in base a particolari fattori quali: la dimensione dell'operatore, il rischio del sistema, il contesto d'uso. In terzo luogo, i principi di sussidiarietà e di proporzionalità chiamano in causa anche la salvaguardia degli spazi di autonomia: la regolamentazione, in pratica, non deve soffocare la sperimentazione e l'innovazione dal basso, ma deve piuttosto creare *frameworks* flessibili che permettano adattamenti.

Nel contesto specifico del diritto d'autore e dell'IA, la proporzionalità significherebbe evitare di estendere meccanicamente ai sistemi generativi tutte quelle tutele che sono state previste per la creazione umana tradizionale. Se, infatti, l'obiettivo è quello di incentivare l'innovazione salvaguardando gli autori umani, forme di protezione *sui generis* calibrate risultano più proporzionate, sia del riconoscimento pieno di autorialità all'IA, sia del divieto assoluto di utilizzo delle opere per l'addestramento. Analogamente, il richiamo alla sussidiarietà suggerirebbe di lasciare spazi per soluzioni contrattuali, *licensing* volontario e pratiche settoriali prima di imporre regimi legislativi rigidi.

*Il bilanciamento* – Un principio cardine della *governance* tecnologica è quello del bilanciamento tra la promozione dell'innovazione e la tutela dei diritti esistenti. Questo principio riconosce che tanto l'innovazione quanto la protezione dei diritti costituiscono valori importanti e che l'adozione di politiche ottimali implica il raggiungimento di compromessi, anziché la massimizzazione di un valore a scapito dell'altro. Nello specifico contesto dell'IA e del diritto d'autore, occorre quindi bilanciare: intanto, l'interesse pubblico al progresso tecnologico contro diritti patrimoniali e morali degli autori; secondariamente, l'accesso universale alla conoscenza contro incentivi economici alla creazione; non da ultimo, la libertà di ricerca contro la tutela della proprietà intellettuale.

Vien da sé che il bilanciamento non si traduce in un'operazione meramente tecnica, ma implica tutta una serie di scelte valoriali fatte a monte. Ed è in tal senso che, diversi approcci e diverse teorie normative possono condurre a bilanciamenti dal risultato assolutamente differente. Ad esempio: l'utilitarismo massimizzerebbe il benessere aggregato anche sacrificando i diritti individuali; il libertarismo proteggerebbe rigorosamente i diritti di proprietà anche a costo di ostacolare l'innovazione; gli approcci comunitaristi privilegierebbero i beni collettivi come il patrimonio culturale comune.

Nelle more del confronto con l'IA generativa, i criteri procedurali che devono ispirare il bilanciamento affinché possa risultare giustificabile sono: la trasparenza nelle scelte valoriali (si tratta, cioè, di rendere esplicito quali

valori si privilegiano e perché); l'inclusività nel processo deliberativo (che significa coinvolgere gli *stakeholder*); la revisione periodica (ovvero, riconoscere che i bilanciamenti ottimali cambiano al mutare delle circostanze tecnologiche ed economiche); la coerenza sistematica (ossia, evitare contraddizioni tra i diversi ambiti normativi)<sup>25</sup>. Chiaramente, tali criteri non determinano univocamente il contenuto delle politiche, ma comunque sia ne disciplinano il processo di elaborazione e giustificazione.

### 3. *Strumenti normativi innovativi*

La regolamentazione tradizionale dell'innovazione tecnologica si è storicamente basata su strumenti quali le leggi, i regolamenti amministrativi e gli interventi giurisprudenziali. Tuttavia, la rapidità dell'evoluzione tecnologica e la complessità dei sistemi di IA richiedono l'integrazione di tali strumenti tradizionali con approcci innovativi più flessibili e adattivi<sup>26</sup>. Di qui, l'utilità di prendere in considerazione tre differenti categorie di strumenti che stanno emergendo nella *governance* dell'IA: *i) Soft law* e co-regolamentazione (che combinano norme volontarie e disposizioni vincolanti); *ii) Regulatory sandboxes* (che permettono sperimentazione controllata); *iii) Standardizzazione tecnica* (che opera come forma di regolamentazione indiretta attraverso definizione di specifiche tecniche).

*i) Soft law e co-regolamentazione* – quando si parla di *soft law*, in via generale, si fa riferimento a tutti quegli strumenti normativi che – ancorché privi di forza cogente – sono comunque sia dotati di funzione

<sup>25</sup> Sul tema dell'opacità dei processi decisionali automatizzati e della difficoltà di accertare dall'esterno le ragioni del responso algoritmico ai fini del controllo antidiscriminatorio, si veda G. GOMETZ, "Intelligenza artificiale, profilazione e nuove forme di discriminazione", *cit.* il quale evidenzia come tale opacità renda particolarmente arduo provare l'impiego di una caratteristica protetta come fattore determinante del trattamento svantaggioso, pur escludendo che essa costituisca di per sé motivo sufficiente per vietare l'impiego dell'IA a supporto delle decisioni giuridiche.

<sup>26</sup> L'esigenza di integrare gli strumenti regolativi tradizionali con approcci più flessibili e adattivi – capaci di rispondere alla rapidità dell'evoluzione tecnologica senza sacrificare la certezza del diritto – è tematizzata, nella prospettiva della filosofia del diritto italiana, in: F. D'AGOSTINO, *Filosofia del diritto*, Torino, Giappichelli, 2000, In particolare al cap. VII; S. COTTA, *Il diritto nell'esistenza*, *cit.*; A.C. AMATO MANGIAMELI, *Informatica giuridica. Appunti e materiali ad uso di lezioni*, *cit.* In ambito internazionale, sul tema della regolamentazione adattiva delle tecnologie emergenti si veda R. BALDWIN, M. CAVE, M. LODGE, *Understanding Regulation: Theory, Strategy, and Practice*, Oxford, Oxford University Press, 2012.

orientativa. Si pensi, ad esempio, ai codici di condotta, alle linee guida o, ancora, alle raccomandazioni. Più in particolare, poi, nel contesto della governance dell'IA generativa, la nozione di *soft law* si intreccia in modo specifico con le forme di co-regolazione previste dall'*AI Act*, che, all'art. 56, attribuisce all'AI Office il compito di incoraggiare e facilitare l'elaborazione, a livello dell'Unione, di codici di pratica per i modelli GPAI, destinati a contribuire alla corretta applicazione degli obblighi previsti dagli artt. 53 e 55. Si tratta, dunque, di uno strumento di dettaglio tecnico e di coordinamento operativo, che si colloca entro la cornice normativa fissata dal legislatore europeo, senza sostituirsi ad essa<sup>27</sup>.

L'elemento che contraddistingue la co-regolamentazione è costituito dalla capacità di combinare elementi di regolamentazione pubblica e di *self-regulation* privata. In pratica, si può dire che lo Stato stabilisce obiettivi e principi generali attraverso Leggi-quadro, delegando, poi, al settore privato (singole imprese, associazioni industriali) l'elaborazione di norme più dettagliate e di meccanismi di implementazione. Tale approccio sfrutta le competenze tecniche del settore e la sua capacità di adattamento rapido, mantenendo l'egida di una supervisione pubblica finalizzata a garantire che gli interessi generali siano comunque garantiti e rispettati.

Nel contesto dell'IA, fra i più noti esempi di *soft law* e di co-regolamentazione possono ricordarsi: i *Principi etici sull'IA* dell'OCSE (adottati da 42 paesi); le *Guidelines* della Commissione Europea per un'IA affidabile; i Codici di condotta volontari promossi da associazioni industriali come Partnership on AI. L'*AI Act*, da un lato, fa ampio ricorso agli standard armonizzati e, in mancanza, alle specifiche comuni ai fini della dimostrazione della conformità; dall'altro lato, distingue i codici di pratica per i modelli GPAI (art. 56) dai codici di condotta previsti, in via volontaria, per sistemi diversi da quelli ad alto rischio (art. 95)<sup>28</sup>.

<sup>27</sup> L'art. 56 dell'*AI Act* prevede che l'*AI Office* incoraggi e faciliti l'elaborazione, a livello dell'Unione, di codici di pratica per i modelli di IA per finalità generali, destinati a contribuire alla corretta applicazione degli obblighi previsti dagli artt. 53 e 55. Tali codici sono elaborati con il coinvolgimento dei *provider*, delle autorità nazionali competenti e degli altri portatori di interessi pertinenti, inclusi rappresentanti della società civile, dell'industria, del mondo accademico e delle PMI. I *provider* dei modelli GPAI con rischio sistemico possono avvalersi dei codici di pratica per dimostrare la conformità agli obblighi di cui all'art. 55 fino alla pubblicazione di uno *standard* armonizzato; ciò non equivale, tuttavia, a una presunzione automatica di conformità. Quanto ai modelli con rischio sistemico, l'art. 51, par. 2, presume *high-impact capabilities* quando l'ammontare cumulativo di calcolo utilizzato per l'addestramento del modello sia superiore a  $10^{25}$  *floating point operations*.

<sup>28</sup> L'*AI Act* prevede due distinti meccanismi di rinvio a *standard* tecnici. In primo luogo, l'art. 40 stabilisce che i sistemi di IA conformi a *standard* armonizzati, o a parti di

Al di là delle necessità – e pure dei vantaggi – bisogna ammettere che *soft law* e co-regolamentazione presentano dei limiti intrinseci. Basti pensare che l'assenza di sanzioni per violazioni può ridurre l'*effectiveness*: in questo senso, le imprese potrebbero aderire formalmente ai codici senza tuttavia implementarli sostanzialmente. Si aggiunga che il rischio di *regulatory capture* è elevato: associazioni industriali potrebbero elaborare norme deboli che servono interessi settoriali piuttosto che pubblici. Non da ultimo, un aspetto di possibile criticità è costituito dal fatto che la molteplicità di iniziative volontarie fra loro non sempre coordinate può generare una notevole frammentazione piuttosto che una situazione di convergenza. Per queste ragioni, l'idea di ricorrere a meccanismi di *soft law* si rivela più efficace come strumento complementare anziché come via alternativa e sostitutiva dell'*hard law* vincolante.

ii) *Regulatory sandboxes per la sperimentazione controllata* – Nel dettaglio, quando si parla di *regulatory sandboxes*, si fa riferimento ad ambienti controllati al cui interno le imprese possono testare prodotti, servizi o modelli di *business* innovativi, beneficiando di esenzioni temporanee dal rispetto di alcune norme regolatorie e agendo sotto la supervisione dell'autorità competente. Originariamente sviluppati nel settore *fintech* dal Regno Unito, i *sandboxes* si sono diffusi in vari paesi e nell'ambito di diversi settori, come utile strumento in grado di bilanciare la spinta verso l'innovazione e la necessità di assicurare un'adeguata protezione dei consumatori. Nel contesto dell'IA, un *sandbox* permetterebbe, ad esempio, di testare sistemi generativi con funzionalità inedite e all'avanguardia in contesti circoscritti (numero limitato di utenti, settore specifico, durata determinata), con il vantaggio di non dover soddisfare tutti quei requisiti che sono previsti dalla normativa ordinaria. In modo particolare, durante la fase di *test*, l'autorità monitora attentamente il funzionamento, raccoglie dati su rischi e benefici, e può intervenire nel caso in cui emergano problemi. Al termine del test, se

essi, i cui riferimenti siano stati pubblicati nella Gazzetta ufficiale dell'Unione europea, beneficano di una presunzione di conformità ai requisiti del Regolamento coperti da tali *standard*. In secondo luogo, l'art. 41 consente alla Commissione di adottare specifiche comuni qualora non esistano *standard* armonizzati adeguati o qualora questi non soddisfino sufficientemente i requisiti applicabili. Distinta da tale meccanismo è la già citata disciplina dei codici di pratica per i modelli GPAI di cui all'art. 56 e dei codici di condotta previsti, in via volontaria, dall'art. 95 per sistemi diversi da quelli ad alto rischio. A livello internazionale, il processo di normazione tecnica in materia di IA è sviluppato, fra gli altri, da ISO/IEC JTC 1/SC 42, cui si devono *standard* quali ISO/IEC 42001:2023 e ISO/IEC 23894:2023. La convergenza tra normazione tecnica e regolamentazione giuridica costituisce così uno dei pilastri dell'approccio europeo alla *governance* dell'IA, in quanto consente di aggiornare i requisiti tecnici specifici senza ricorrere continuamente alla revisione legislativa.

i risultati sono positivi, il sistema può essere autorizzato per uso generalizzato; laddove, invece, gli esiti siano negativi, il modello sarà ritirato. A ciò si aggiunga che le conoscenze acquisite si rivelano dirimenti (e strategiche) per introdurre opportuni aggiustamenti normativi.

I *sandboxes* presentano diversi vantaggi: riducono le barriere normative all'ingresso per *startup* e innovatori; permettono apprendimento regolativo *evidence-based*; facilitano il dialogo costruttivo tra regolatori e regolati; generano flessibilità in contesti di rapida evoluzione tecnologica. L'*AI Act* prevede esplicitamente il ricorso a *regulatory sandboxes* per l'IA (l'art. 57 stabilisce l'obbligo per le autorità nazionali di istituirli almeno a livello nazionale entro il 2 agosto 2026; l'art. 58 ne definisce le condizioni di accesso, prevedendo un regime di priorità per PMI e *startup*; l'art. 59 disciplina le responsabilità delle autorità competenti nella supervisione; l'art. 60 regola il trattamento di dati personali nelle attività di sviluppo e test)<sup>29</sup>.

Di qui, però, anche tutta una serie di possibili questioni: intanto, il possibile rischio di cattura (le imprese partecipanti potrebbero influenzare impropriamente l'autorità); si aggiunga, l'eventuale limitazione nell'equità dell'accesso (se solo alcune imprese possono partecipare, ciò crea vantaggi competitivi ingiustificati); ancora, alcune criticità circa la protezione dei partecipanti (chi testa sistemi sperimentali assume rischi; occorre garantire consenso informato e meccanismi di compensazione per danni); e, non da ultimo, anche questioni in ordine alla transferibilità dei risultati (quanto appreso in *sandbox* può non generalizzarsi a contesti reali più complessi).

*iii) Standardizzazione tecnica come forma di regolamentazione indiretta* – La standardizzazione tecnica consiste nell'elaborazione di specifiche

<sup>29</sup> I *regulatory sandboxes* per l'IA sono disciplinati dagli artt. 57-63 del Reg. (UE) 2024/1689 (*AI Act*). In particolare: l'art. 57 impone agli Stati membri di istituire almeno un *sandbox* nazionale entro il 2 agosto 2026; l'art. 58 ne definisce le condizioni di accesso, prevedendo un regime di priorità per PMI e *startup*; l'art. 59 stabilisce le responsabilità delle autorità competenti nella supervisione delle attività svolte nel *sandbox*; l'art. 60 disciplina in modo specifico il trattamento di dati personali nelle attività di sviluppo e test condotte all'interno del *sandbox*. Il meccanismo consente alle imprese di sperimentare sistemi di IA innovativi in un ambiente controllato, con deroghe temporanee e circoscritte a talune disposizioni normative applicabili, sotto la supervisione dell'autorità competente. Al termine del periodo di sperimentazione, i risultati e le conoscenze acquisite sono messi a disposizione dell'autorità e, in forma anonimizzata, delle istituzioni europee, a fini di apprendimento regolatorio. Il modello dei *sandboxes* è stato originariamente sviluppato dalla *Financial Conduct Authority* britannica nel settore *fintech* (2016) e successivamente adottato da numerose autorità di regolamentazione in tutto il mondo.

tecniche dettagliate che definiscono requisiti, metodi di test, procedure operative per prodotti o servizi. Gli *standard* possono essere *de jure* (ossia, adottati da organismi ufficiali come ISO, CEN) o *de facto* (vale a dire, emergenti dal mercato). Pur essendo formalmente volontari, gli *standard* operano come forma di regolamentazione indiretta: diventano condizione per accesso al mercato, sono incorporati in contratti, possono essere resi vincolanti per riferimento normativo.

Nel settore dell'IA, la standardizzazione sta emergendo come strumento chiave di *governance*. Organizzazioni come ISO hanno sviluppato *standard* su aspetti quali gestione del rischio<sup>30</sup>, *governance* organizzativa dell'IA<sup>31</sup> e panoramica sulla *trustworthiness* dei sistemi AI, che include la robustezza<sup>32</sup> pur non riducendosi ad essa. L'*AI Act* fa ampio riferimento a *standard* armonizzati, affermando, peraltro, che i fornitori che conformano i propri sistemi a tali *standard* beneficiano di una presunzione di conformità ai requisiti essenziali del Regolamento stesso.

Ora, come è evidente, la standardizzazione, per un verso, presenta dei vantaggi notevoli (incorpora *expertise* tecnica specializzata, facilita interoperabilità tra sistemi diversi; fornisce certezza giuridica sulle modalità di *compliance*, evolve più rapidamente della legislazione), per un altro verso, tuttavia, essa solleva anche svariate preoccupazioni. Il processo di standardizzazione, infatti, è dominato da grandi imprese e dai paesi più sviluppati che hanno le risorse necessarie per partecipare; di contro, le piccole imprese e i paesi in via di sviluppo rischiano l'esclusione. Gli *standard* possono consolidare approcci tecnici specifici scoraggiando innovazione radicale. Inoltre, la tecnicità rende difficile per gli *stakeholder* non esperti (consumatori, organizzazioni sociali) partecipare efficacemente.

<sup>30</sup> ISO/IEC, ISO/IEC 23894:2023 *Information technology – Artificial intelligence – Guidance on risk management*, Ginevra: International Organization for Standardization, 2023.

<sup>31</sup> ISO/IEC, ISO/IEC 38507:2022 *Information technology – Governance of IT – Governance implications of the use of artificial intelligence by organizations*, Ginevra: International Organization for Standardization, 2022

<sup>32</sup> ISO/IEC, ISO/IEC TR 24028:2020 *Information technology – Artificial intelligence – Overview of trustworthiness in artificial intelligence*, Ginevra: International Organization for Standardization, 2020.

## CAPITOLO VIII

### SCENARI FUTURI E PROSPETTIVE DI RICERCA

SOMMARIO: 1. Evoluzioni tecnologiche prevedibili e loro impatto giuridico. – 2. Ripensare i fondamenti del diritto d'autore. – 3. L'intelligenza umana e quella artificiale: confini e relazioni.

#### 1. *Evoluzioni tecnologiche prevedibili e loro impatto giuridico*

Come si è detto, quello dell'IA è un campo in evoluzione rapidissima, ragion per cui qualsiasi analisi che si limitasse a scrutare il presente rischierebbe di incorrere in un'obsolescenza prematura. Il giurista (teorico e pratico) è dunque chiamato a rivolgere lo sguardo al futuro, sforzandosi di interrogarsi su quali sviluppi tecnologici saranno prevedibili nel medio-lungo termine e quali ulteriori sfide si paleseranno al cospetto del diritto, della filosofia e dell'etica.

Certo, prevedere il futuro tecnologico è un esercizio intrinsecamente incerto. Del resto, nella storia della tecnologia, molte previsioni si sono rivelate clamorosamente errate: si sottovalutarono le potenzialità del *computer* personale, di Internet, degli *smartphone*; si sopravvalutarono promesse di tecnologie che non si sono mai materializzate. Ciò detto, l'incertezza non giustifica – né può di per sé giustificare – l'astensione dalla riflessione prospettica. Al contrario, proprio perché il futuro è aperto e dipende dalle scelte che compiamo, occorre anticipare gli scenari possibili per orientare consapevolmente lo sviluppo tecnologico verso esiti desiderabili.

Ora, giunti a questo punto del nostro percorso, è interessante volgere l'attenzione a tre possibili traiettorie tecnologiche che – sebbene ancora incerte quanto a tempi e modalità di realizzazione – appaiono sufficientemente plausibili da meritare una certa attenzione. Si tratta: 1) del progresso verso l'IA generale (AGI), capace di *performance* cognitive comparabili o superiori a quelle umane su qualsiasi compito; 2) dell'emergere dell'IA quantistica, che sfrutterebbe principi di meccanica quantistica per capacità computazionali radicalmente superiori; 3) dello sviluppo di interfacce cervello-computer che permetterebbero ibridazione diretta tra intelligenza biologica e artificiale. Prospettive, che sollevano interrogativi filosofici e giuridici ancor più complessi e ostici di quelli suscitati dall'odierna IA generativa.

1) Si osservi, l'IA attuale, per quanto strabiliante e impressionante, rimane *Narrow AI*, ossia IA ristretta: si tratta, invero, pur sempre di sistemi specializzati che sono eccellenti nella realizzazione di compiti specifici ma che si rivelano incapaci di generalizzare oltre il dominio per il quale sono stati addestrati. Un sistema che gioca a scacchi a livello mondiale non sa guidare un'automobile; un modello che genera immagini fotorealistiche non comprende il linguaggio naturale. L'IA generale (AGI), invece, designa sistemi ipotetici dotati di capacità cognitive flessibili e generali comparabili a quelle umane. Vale a dire: capacità di apprendimento trasferibile tra domini diversi, attitudine al ragionamento astratto, comprensione causale profonda, autonomia negli obiettivi. La previsione circa l'*an* e il *quando* dell'AGI è oggetto di disaccordo.

Alcuni ricercatori, come Ray Kurzweil<sup>1</sup>, formulano previsioni molto ottimistiche. In *The Singularity Is Nearer*, Kurzweil riprende e aggiorna la propria previsione, già formulata alla fine degli anni Novanta, secondo cui l'intelligenza artificiale raggiungerà un livello umano entro il 2029, mentre la singolarità tecnologica si collocherebbe intorno al 2045, come esito del progresso esponenziale della capacità computazionale e del perfezionamento degli algoritmi. Altri, invece, come Gary Marcus<sup>2</sup>, sostengono che l'approccio attuale basato sul *deep learning* presenti limiti fondamentali e che l'AGI richieda *breakthrough* concettuali non ancora all'orizzonte. Secondo la *AI Impacts Survey 2022*<sup>3</sup> – condotta su 738

<sup>1</sup> Tra gli scritti più recenti e coerenti con l'affermazione testuale richiamata nel testo circa il raggiungimento dell'AGI entro il 2029 e della singolarità tecnologica entro il 2045, si veda, per tutti, R. KURZWEIL, *The Singularity Is Nearer: When We Merge with AI*, New York, Viking, 2024; ID., *The Singularity Is Near: When Humans Transcend Biology*, New York, Viking, 2005. La differenza principale tra le due opere sta nel grado di dettaglio e nei dati a supporto: nel volume del 2024, KURZWEIL utilizza i progressi tecnologici concreti avvenuti tra il 2005 e il 2024 per rinforzare le sue tesi originali, confermando che, secondo lui, il percorso verso l'AGI e la Singolarità non ha subito ritardi sostanziali. In sintesi, non ha cambiato idea sulle scadenze, ma ha riaffermato con forza che i tempi sono maturi.

<sup>2</sup> Per sostenere la tesi dei limiti strutturali degli approcci attuali fondati sul *deep learning* e della necessità di *breakthrough* concettuali ulteriori, cfr. G. MARCUS, E. DAVIS, *Rebooting AI: Building Artificial Intelligence We Can Trust*, New York, Pantheon Books, 2019.

<sup>3</sup> K. GRACE, J. SALVATIER, A. DAFOE, B. ZHANG, O. EVANS, "Viewpoint: When Will AI Exceed Human Performance? Evidence from AI Experts", in *Journal of Artificial Intelligence Research*, 62, 2018, pp. 729-754; cfr. anche K. GRACE, J. SELBST, S. BEARD, O. SANDBERG, *AI Impacts Survey 2022: Expert Survey on Progress in AI*, Berkeley, Machine Intelligence Research Institute (MIRI), 2022. Il primo studio, del 2018, definisce il quadro metodologico fondamentale per valutare le stime temporali sul progresso verso l'intelligenza artificiale di alto livello; il secondo aggiorna tale valutazione, riflettendo le accelerazioni occorse nel dibattito scientifico tra il 2018 e il 2022.

risposte, alcune parziali, di ricercatori che avevano pubblicato a *NeurIPS* o *ICML* nel 2021 – il 50% *aggregate forecast* per il raggiungimento della *High-Level Machine Intelligence* (concetto affine ma non identico all'AGI in senso filosofico pieno) cade nel 2059, con una dispersione molto ampia delle stime individuali.

Indipendentemente dalla tempistica ipotizzata e ipotizzabile, resta il fatto – incontrovertibile – che l'AGI pone interrogativi filosofici (prima ancora che giuridici) radicali. Se un sistema artificiale possedesse tutte le capacità cognitive umane (ragionamento, comprensione, autoconsapevolezza, forse anche coscienza) potrebbe ancora negarsi che esso sia un agente morale meritevole di considerazione etica? I criteri tradizionali di personalità basati su capacità razionali, autonomia e intenzionalità verrebbero soddisfatti, o, al contrario, l'intuizione secondo la quale un artefatto non potrà mai essere equiparato a un essere umano rimarrebbe nonostante tutto ugualmente forte. E, ancora, come si potrebbero bilanciare considerazioni funzionaliste (ciò che conta è che cosa un sistema può fare) con quelle essenzialiste (ciò che conta è che cosa un sistema è per natura)?

Nel contesto specifico del diritto d'autore, poi, l'AGI metterebbe definitivamente in crisi il presupposto antropocentrico. Se un'AGI potesse creare opere letterarie, artistiche, musicali indistinguibili da quelle umane non solo formalmente ma anche quanto a profondità espressiva e significato, sarebbe ancora sostenibile negare protezione a tali opere? E se le si concedesse protezione, con quale giustificazione? Bisognerebbe riprendere la teoria lockiana del lavoro, interrogandosi se l'AGI possa propriamente dirsi 'lavoratrice'? La teoria hegeliana della personalità, verificando se l'AGI possa essere considerata dotata di una personalità rilevante? Oppure quella utilitaristica, che forse rimane l'unica teoria pienamente applicabile?

2) Si badi, mentre i *computer* attuali rappresentano informazioni come *bit* che assumono valore zero o uno, i *qubit* dei computer quantistici<sup>4</sup> possono esistere in sovrapposizione di stati e sfruttare l'*entanglement* e l'interferenza quantistica per risolvere in modo efficiente classi specifiche di problemi (come la fattorizzazione di numeri interi o la simulazione di sistemi quantistici) che risulterebbero intrattabili su *hardware* classico. Ora, sebbene l'informatica quantistica sia ancora in una fase relativamente precoce di sviluppo e i sistemi esistenti restino limitati e fragili,

<sup>4</sup> L'informatica quantistica sfrutta principi della meccanica quantistica – sovrapposizione, *entanglement* – per eseguire calcoli che sarebbero impraticabili su computer classici. Per una introduzione d'insieme, cfr. M. BROOKS, *Quantum Computing: Progress and Prospects*, Washington DC, National Academies Press, 2019.

progressi significativi sono in corso e si registrano prime forme di *deployment* e di applicazione concreta nel più ampio settore delle tecnologie quantistiche, pur restando lontana una piena maturità dell'informatica quantistica generalista.

In modo particolare, l'integrazione tra *machine learning* e computazione quantistica potrebbe generare l'IA quantistica. Vale a dire, sistemi di IA che sfruttano vantaggi computazionali quantistici per capacità di apprendimento e generazione superiori. Ricerche teoriche suggeriscono che algoritmi quantistici potrebbero accelerare drammaticamente l'addestramento di reti neurali, esplorare spazi di soluzione inaccessibili classicamente e generare *output* con proprietà emergenti imprevedibili. Tuttavia, la realizzazione pratica di un'IA quantistica scalabile rimane una sfida aperta e alcuni ricercatori dubitano che i vantaggi teorici si traducano in applicazioni concrete<sup>5</sup>.

Se l'IA quantistica si materializzasse, le sue capacità creative potrebbero trascendere quelle di qualsiasi IA oggi immaginabile. Sistemi quantistici potrebbero generare opere d'arte basate su *pattern* che riflettono proprietà fondamentali della realtà fisica a livello quantistico, inaccessibili alla percezione e all'intuizione umana. La musica, la pittura, la letteratura prodotte potrebbero incorporare strutture di complessità tale da sfuggire completamente alla comprensione umana, pur essendo esteticamente apprezzabili. In tale scenario, la distinzione stessa tra creazione umana e artificiale perderebbe ogni significato: l'IA quantistica sarebbe genuinamente aliena, operante in spazi concettuali radicalmente diversi da quelli umani.

Vien da sé che le implicazioni giuridiche di un siffatto scenario sarebbero profondissime oltre che radicali. Il diritto d'autore attuale presuppone che le opere, pur originali, siano comunque comprensibili e fruibili dagli esseri umani. Ma come tutelare opere che trascendono la comprensibilità umana? Chi potrebbe valutare se un'opera quantistica è originale o derivativa? E soprattutto, quale valore avrebbero tali opere per una so-

<sup>5</sup> Lo scetticismo circa la traduzione pratica dei vantaggi teorici dell'informatica quantistica in applicazioni di intelligenza artificiale scalabili è argomentato in F. ARUTE et al., "Quantum Supremacy Using a Programmable Superconducting Processor", in *Nature*, 574, 2019, pp. 505-510, che pur dimostrando la supremazia quantistica su compiti specifici segnala i limiti della generalizzazione; cfr. anche S. AARONSON, *Quantum Computing Since Democritus*, Cambridge, Cambridge University Press, 2013 e M. BROOKS, *Quantum Computing: Progress and Prospects*, cit. Sul piano più specificamente filosofico-giuridico, il problema della comprensibilità umana degli *output* di sistemi che operano in spazi computazionali inaccessibili alla percezione ordinaria pone interrogativi radicali circa i criteri di valutazione dell'originalità e della tutela autorale, che il diritto positivo attuale non è attrezzato ad affrontare.

cietà composta da esseri umani non quantistici? Questi interrogativi, per quanto speculativi, evidenziano che l'evoluzione tecnologica potrebbe non solo richiedere aggiustamenti normativi ma mettere in discussione i presupposti stessi sui quali il diritto è edificato.

3) Le interfacce cervello-computer (*Brain-Computer Interfaces* – BCI) permettono una comunicazione diretta tra il cervello biologico e i dispositivi elettronici, bypassando i canali sensoriali e motori tradizionali. Applicazioni attuali includono protesi neurali che restituiscono funzionalità motorie a persone paralizzate e dispositivi che permettono controllo di *computer* attraverso segnali cerebrali. Tuttavia, la ricerca sta esplorando possibilità ancor più strabilianti e radicali: *enhancement* cognitivo che amplifica memoria, attenzione, velocità di elaborazione; comunicazione diretta cervello-a-cervello; integrazione con IA per *augmented intelligence* in cui le capacità biologiche e quelle artificiali si fondono in un tutt'uno.

Imprese come Neuralink hanno già effettuato i primi impianti su pazienti umani (gennaio 2024, dispositivo N1 con 1024 elettrodi)<sup>6</sup> e dichiarano come obiettivo di lungo periodo l'interfacciamento di milioni di neuroni simultaneamente ad alta banda. Sebbene restino forti ambizioni di sviluppo verso interfacce a banda sempre più elevata, il quadro tecnico attuale rimane ancora sperimentale e lontano da una piena maturità clinica e applicativa. Questa prospettiva pone questioni esistenziali inerenti alla ridefinizione dell'identità umana<sup>7</sup>, poiché il raggiungimento di tali livelli di *augmentation* mette in discussione la persistenza dei tratti fondamentali della nostra specie. Emerge dunque la necessità di tracciare un confine netto tra l'*enhancement* accettabile, inteso come potenziamento

<sup>6</sup> Il primo impianto del dispositivo N1 di Neuralink su un paziente umano è avvenuto il 29 gennaio 2024. Il dispositivo, denominato "Telepathy", è composto da 1024 elettrodi distribuiti su 64 fili sottilissimi impiantati nella corteccia motoria e consente al paziente di controllare dispositivi digitali attraverso segnali cerebrali. Cfr. Neuralink, *First Human Receives Neuralink Implant*, comunicato stampa, 30 gennaio 2024, disponibile su neuralink.com. Per una riflessione filosofico-giuridica sull'integrazione tra menti umane, corpi e intelligenza artificiale, cfr. A.C. AMATO MANGIAMELI, "Intelligenza artificiale generativa e Metaverso. A proposito di menti e corpi fusi con IA", in *Rivista di Filosofia del Diritto*, 1/2024, pp. 45-55.

<sup>7</sup> Cfr.; M. GAWDAT, *Superintelligenti. Come salvare il mondo dall'intelligenza artificiale*, trad. it., Milano, Rizzoli, 2022; J. Le Doux, *Lunga storia di noi stessi, Come il cervello è diventato cosciente*, trad. it., Milano, Cortina, 2020; J. LOVELOCK, B. APPEYARD, *Novacene. L'era dell'iperintelligenza*, trad. it., Torino, Bollati Boringhieri, 2020; S. AMATO, "Darwin tra le macchine? L'intelligenza della tecnoscienza", in *Teoria e Critica della Regolazione Sociale*, 2/2022, pp. 13-30; ID., "Macchinazioni. Soggettività, agentività ed entità esotiche", in *L'Ircocervo*, 1/2024, pp. 289-306.

delle facoltà esistenti, e una trasformazione radicale che rischierebbe di evolvere in una condizione post-umana, strutturalmente diversa da quella biologica originaria. Del resto, il superamento di determinati limiti cognitivi e fisici sposta inevitabilmente la demarcazione tra l'integrazione tecnologica e la mutazione della specie, imponendo una riflessione sulla perdita della natura umana in favore di una nuova identità sintetica.

Ovviamente, nel contesto creativo, le BCI potrebbero rivoluzionare integralmente il processo artistico. Un musicista potrebbe comporre sinfonie traducendo direttamente immagini mentali in suono senza mediazione fisica; un pittore potrebbe generare opere visualizzando mentalmente forme e colori che l'IA trasmette istantaneamente su tela; uno scrittore potrebbe produrre romanzi alla velocità del pensiero. In tale scenario, la distinzione tra autore umano e IA si dissolverebbe: l'opera nascerebbe da una collaborazione simbiotica in cui è impossibile separare il contributo biologico e quello artificiale.

Le implicazioni giuridiche sarebbero vertiginose. A chi spetterebbero i diritti su opere create attraverso le BCI: all'essere umano-*cyborg*, all'IA integrata, oppure a entrambi? Come sarebbe possibile distinguere tra idee spontanee e suggerimenti algoritmici quando la comunicazione è diretta e inconscia? I diritti morali avrebbero ancora senso quando l'identità personale è parzialmente costituita da componenti artificiali sostituibili?

## 2. *Ripensare i fondamenti del diritto d'autore*

Gli scenari tecnologici ipotizzati nel paragrafo precedente evidenziano che le sfide poste dall'IA al diritto in generale e *a fortiori* al diritto d'autore potrebbero essere solo l'inizio di una serie di trasformazioni ancora più radicali. Di fronte a tale prospettiva, occorre interrogarsi se il diritto d'autore moderno – che, come si è avuto modo di sottolineare, è edificato su concetti elaborati tra il XVIII e il XIX secolo – rappresenti ancora un *framework* adeguato a regolare la creatività o se sia necessario un suo completo ripensamento, in vista dell'elaborazione di nuovi paradigmi.

Come s'è detto, il diritto d'autore moderno è profondamente antropocentrico: presuppone, cioè, che solo esseri umani possano essere riconosciuti quali autori di un'opera e che la creatività sia manifestazione esclusiva di facoltà specificamente umane. Di qui, l'idea che il valore delle opere derivi dal loro legame con la personalità umana che le genera. È evidente che l'IA – e ancor di più le sue prospettive future – mettono in crisi questo radicale antropocentrismo del diritto d'autore, ponendo l'accento sulla possibilità (o meno) di un suo superamento. In partico-

lare, si fa strada l'idea di un diritto d'autore post-antropocentrico, che riconoscerebbe la possibilità che la creatività emerga anche da processi non umani e che persino artefatti artificiali possono meritare stessa tutela giuridica che viene – da sempre – riconosciuta in capo all'autore-artista-creativo-persona. Ovviamente si fanno strada differenti ipotesi e diverse varianti del diritto d'autore post-antropocentrico.

Una prima variante, che potremmo definire moderata, manterrebbe la centralità dell'autore umano ma riconoscerebbe anche forme di tutela complementare per le produzioni artificiali: il *copyright* pieno per opere umane, protezione *sui generis* ridotta per le opere dell'IA, regime intermedio per le creazioni ibride. Questa soluzione preserva la distinzione fra umano e artificiale, gerarchizzando i valori: la creatività umana rimane superiore e meritevole di maggiore protezione, ciononostante si riconosce che anche la creatività artificiale ha un suo valore.

Una seconda variante, di stampo più radicale, abbandonerebbe il criterio dell'autorialità umana come determinante e valuterebbe invece le opere esclusivamente in base alle loro caratteristiche intrinseche: originalità, complessità, valore estetico o sociale. In questa prospettiva, non importa affatto se un'opera è stata creata da un essere umano, da un'IA o da collaborazione tra i due; ciò che conta è che l'opera soddisfi i requisiti di protezione valutati oggettivamente. Questa impostazione riflette una concezione funzionalista radicale e comporterebbe riscrittura completa del diritto d'autore.

Un altro dibattito – comunque parallelo a quello sul diritto d'autore e alla sua impostazione antropocentrica – concerne la proprietà intellettuale. Due posizioni radicali si confrontano: quella abolizionista, che sostiene l'eliminazione del diritto d'autore e dei brevetti; e quella massimalista, che propone un rafforzamento ed un'estensione della tutela. Entrambe le posizioni sono sorrette da argomenti filosofici ed economici degni di nota ed entrambe potrebbero trovare nuova rilevanza nell'era dell'IA.

Nel dettaglio, la posizione abolizionista – difesa da autori come Michele Boldrin e David Levine in *Against Intellectual Monopoly*<sup>8</sup> – sostiene che la proprietà intellettuale crea più danni che benefici. Gli argomenti principali avanzati da questa posizione sono: la conoscenza è bene non rivale (il mio utilizzo non impedisce il tuo), dunque, escluderla è inefficiente; i monopoli temporanei rallentano innovazione impedendo utilizzazione e miglioramento di idee esistenti; evidenze empiriche non dimostrano che paesi con

<sup>8</sup> M. BOLDRIN, D.K. LEVINE, *Against Intellectual Monopoly*, Cambridge, Cambridge University Press, 2008.

protezione più forte innovino di più; modelli alternativi di incentivo (premi, patronage pubblico, produzione *peer-to-peer*) possono sostituire proprietà esclusiva. Nell'era dell'IA, inoltre, l'abolizione sarebbe giustificata ulteriormente. Infatti, se i costi di creazione tendono a zero grazie all'automazione, la funzione incentivante della protezione diventa superflua.

La posizione massimalista, al contrario, sostiene che nell'economia della conoscenza la proprietà intellettuale diventa ancora più cruciale. Gli investimenti in ricerca, sviluppo, creazione sono enormi e rischiosi; ragion per cui, senza la previsione di una protezione solida nessuno investirebbe. L'IA amplifica questo bisogno: sviluppare modelli generativi costa centinaia di milioni di dollari; permettere *free-riding* una volta che il modello esiste eliminerebbe incentivi. Inoltre, i diritti morali degli autori umani richiedono tutela rafforzata proprio per distinguere la creazione autentica dalla produzione meccanica. Questa posizione potrebbe proporre un'estensione della durata della protezione, l'ampliamento dell'ambito e il rafforzamento dell'*enforcement*.

Una posizione intermedia, di stampo pragmatico, riconosce che né l'abolizione né la massimalizzazione sono ottimali. Ciò che – a detta di questa posizione – il diritto d'autore deve fare oggi, difatti, è evolversi bilanciando dinamicamente incentivi, accesso e giustizia distributiva. Nell'era dell'IA, ciò potrebbe significare: una protezione forte per opere umane autentiche, debole o assente per produzioni automatizzate; una durata differenziata; eccezioni ampie per usi trasformativi e ricerca; meccanismi redistributivi che garantiscano accesso universale.

Un'ultima prospettiva, ancora più visionaria, propone di abbandonare completamente il paradigma dei diritti autoriali fondato sull'attribuzione individuale e sul controllo esclusivo, sostituendolo con paradigmi basati su diritti collettivi alla conoscenza e alla cultura. Questa visione si ispira a tradizioni comunitarie e a concezioni della conoscenza come bene comune (*commons*), che deve essere accessibile a tutti e gestito collettivamente.

Il filosofo Michael Hardt e il teorico politico Antonio Negri, in *Commonwealth*<sup>9</sup> hanno sviluppato il concetto di *comune* (*the common*) per designare ricchezze prodotte socialmente, che non possono essere privatizzate senza che ciò arrechi ingiustizia. La conoscenza, le opere culturali, i codici genetici appartengono alla comunità e ricadono, quindi, nella nozione di bene comune. Per questo motivo, appropriarle privatamente (o, meglio, riservarle alla proprietà di qualcuno) significa espropriare rispetto alla col-

<sup>9</sup> M. HARDT, A. NEGRI, *Commonwealth*, Cambridge (MA), Harvard University Press, 2009.

lettività (o, più propriamente, significa sottrarle ad un'idea di condivisione e di patrimonialità condivisa e diffusa in maniera indistinta, egualitaria). Questa prospettiva suggerirebbe che tutte le opere, umane o artificiali, debbano entrare immediatamente nel dominio pubblico, gestite attraverso istituzioni democratiche che ne garantiscano accesso universale<sup>10</sup>.

Un'altra tradizione rilevante è quella dei *Creative Commons*<sup>11</sup> e delle licenze *copyleft*<sup>12</sup>, che hanno dimostrato la possibilità di sviluppare una produzione culturale non proprietaria su scala massiva. Piattaforme come

<sup>10</sup> È interessante ricordare che il termine latino *communis* – da cui derivano tanto “comunione” quanto “comunicazione” – designa originariamente ciò che è condiviso, partecipato, non esclusivo: etimologicamente si compone di *cum* (insieme) e *munus* (dono, dovere, ufficio): il *commune* è quindi il dono che si mette in comune, il vincolo che lega i membri di una comunità attraverso la reciprocità dell'obbligo e del beneficio. Questa radice semantica illumina la proposta di Hardt e Negri: le opere culturali, le conoscenze, i codici sono *communio* non perché siano privi di valore, ma perché il loro valore è intrinsecamente relazionale e si moltiplica nella condivisione anziché diminuire nel consumo. Cfr. anche A.C. AMATO MANGIAMELI, “Sul diritto d'autore e la proprietà intellettuale, ovvero l'opera dell'ingegno tra dono, diritto e tecnica”, in Ead., *Informatica giuridica. Appunti e materiali ad uso di lezioni*, Torino, Giappichelli, 2015, pp. 297-310.

<sup>11</sup> Le licenze Creative Commons costituiscono un sistema di licenze pubbliche standardizzate che consente all'autore di modulare anticipatamente le condizioni di utilizzazione dell'opera, favorendone la circolazione e il riuso entro limiti predeterminati. Per la nozione di *Creative Commons*, si vedano, tra gli altri: L. LESSIG, *The Creative Commons*, in *Montana Law Review*, 65, 2004; T. SIMMONDS, *Common Knowledge? The Rise of Creative Commons Licensing*, in *Legal Information Management*, 10(3), 2010, pp. 162-165; L.A. MAGGIO, K. STRANACK, *Understanding Creative Commons*, in *Academic Medicine*, 95(2), 2020, p. 322. In dottrina italiana, si vedano, per tutti: M. RICOLFI, *License contracts, free software and creative commons*, in *Italian National Reports to the XIXth International Congress of Comparative Law*, Milano, Giuffrè, 2014, pp. 493-530, ove il tema è ricondotto al quadro privatistico dei contratti di licenza; M. RICOLFI, M. VAN EECCHOUD, F. MORANDO, P. TZIAVOS, L. FERRAO, *The “Licensing” of Public Sector Information*, in *Informatica e diritto*, 2011, pp. 129-146, anche con riguardo all'impiego di modelli licenziatari aperti e strumenti *Creative Commons* nel contesto degli *open data*.

<sup>12</sup> Il *copyleft* designa un modello di licenza che, facendo leva sul diritto d'autore, consente l'uso, la modifica e la redistribuzione dell'opera a condizione che le opere derivate siano a loro volta diffuse secondo termini analoghi, così da impedire appropriazioni esclusive di sviluppi successive. Per la nozione di *copyleft*, si vedano, tra gli altri: M. MUSTONEN, *Copyleft—the Economics of Linux and other Open Source Software*, in *Information Economics and Policy*, 15(1), 2003, pp. 99-121; C.W. ZHU, ‘*Copyleft*’ *Reconsidered: Why Software Licensing Jurisprudence Needs Insights from Relational Contract Theory*, in *Social & Legal Studies*, 22(3), 2013. In dottrina italiana, si vedano, per tutti: M. RICOLFI, *License contracts, free software and creative commons*, cit., pp. 493-530, ove il fenomeno è inquadrato nel più ampio tema delle licenze di software libero e dei modelli aperti di circolazione delle opere; sullo sfondo teorico del rapporto tra esclusiva proprietaria e interesse pubblico: M. BORGHI, *Il diritto d'autore tra regime proprietario e*

*Wikipedia*, *software open source* come *Linux* e *repository* scientifici ad accesso aperto dimostrano che i modelli non basati su esclusività possono generare contenuti di altissima qualità. L'IA – permettendo automazione della creazione – potrebbe rendere questi modelli ancora più praticabili: se i costi marginali di produzione sono zero, perché mantenere scarsità artificiale attraverso proprietà esclusiva?

Questa visione si scontra, tuttavia, con la realtà economica e giuridica. Dal punto di vista economico, il rischio concreto è che senza incentivi finanziari, il flusso di capitali verso l'innovazione tecnologica rischierebbe di arrestarsi. Il modello del 'bene comune' inoltre espone il sistema al paradosso del sovrasfruttamento delle risorse collettive (la 'tragedia dei *commons*'). C'è poi il pericolo che la tutela della collettività finisca per eclissare il diritto individuale dell'autore a veder protetta l'originalità e la paternità della propria creazione.

Se così, è evidente che meccanismi quali *licensing* collettivo, gestione cooperativa di *commons* digitali, e forme ibride pubblico-privato possono rappresentare degli importanti passi intermedi. L'obiettivo non è certo quello di realizzare l'utopia dell'abolizione completa della proprietà, ma quello di ricercare un equilibrio più giusto ed efficiente tra l'appropriazione privata e la condivisione pubblica della conoscenza.

### 3. *L'intelligenza umana e quella artificiale: confini e relazioni*

Dopo aver menzionato alcuni degli scenari tecnologici futuri e messo in luce le possibili rifondazioni del diritto d'autore, è inevitabile porsi una domanda radicale, che innerva tutto il dibattito sull'IA e sulle sue applicazioni: quale relazione sussiste (o sussisterà) tra l'intelligenza umana e quella artificiale? Sono fenomeni qualitativamente diversi o quantitativamente graduati? La creatività artificiale sostituirà quella umana o la completerà, divenendo un complemento, un irrinunciabile strumento sussidiario? Esiste qualcosa di irriducibilmente umano che resisterà alla seduzione dell'automazione? Queste domande non hanno solo rilevanza teoretica ma comportano implicazioni pratiche profonde per le scelte normative e sociali.

Un dilemma centrale riguarda se l'IA debba essere intesa come strumento complementare che potenzia capacità umane, oppure come sostituto che eventualmente renderà superfluo il contributo umano in molti

ambiti. La risposta a questa domanda non è determinata tecnologicamente, ma dipende dalle scelte sociali, economiche e politiche. Diverse sono le traiettorie possibili.

Uno scenario di *complementarità* prevede che l'IA rimanga strumento sofisticato nelle mani di creatori umani, di cui amplifica le possibilità espressive pur senza sostituire il giudizio, l'intenzionalità e la sensibilità estetica. In questo scenario – così come l'invenzione del pianoforte non eliminò musicisti ma permise nuove forme musicali – l'IA generativa aprirebbe orizzonti creativi inediti senza sostituire gli artisti. Il valore rimarrebbe insito nella visione umana, nell'originalità concettuale, nella capacità di comunicare esperienza e significato. L'IA gestirebbe più che altro quelli che sono gli aspetti tecnici e meccanici, diversamente la direzione artistica resterebbe esclusivamente umana.

Uno scenario di *sostituzione*, invece, prevede che l'IA raggiunga e superi le capacità creative umane in tutti i domini, rendendo il lavoro creativo umano economicamente obsoleto come il lavoro agricolo manuale fu reso obsoleto dalla meccanizzazione. In tale scenario, la creatività nel campo dell'arte, della letteratura e della musica sarebbe meglio soddisfatta dalle produzioni algoritmiche, indubbiamente più abbondanti, più economiche e all'occorrenza anche più duttili e più personalizzabili rispetto a quelle umane. Se così, i creatori umani sopravviverebbero solo in ristrettissime nicchie di mercato *premium* o come *hobby* non economicamente sostenibile.

Quale scenario prevarrà dipende da molteplici fattori. Tecnologicamente, bisogna vedere se l'IA raggiungerà effettivamente *performance* creative indistinguibili o superiori a quelle umane. Economicamente, è necessario valutare se i modelli di *business* basati sull'IA saranno più profittevoli di quelli basati su creativi umani. Culturalmente, è da vedere se il pubblico continuerà a valorizzare l'autenticità e l'autorialità umana o se preferirà dei contenuti algoritmici. Non da ultimo, politicamente, bisogna capire se le società sceglieranno di proteggere lavoro creativo umano attraverso regolamentazione, sussidi o altri meccanismi.

Al di là di quelle che saranno le evoluzioni, è importante sottolineare un aspetto che non deve essere affatto trascurato, ovvero che anche nello scenario complementare, la creatività sarà profondamente trasformata. In modo particolare, si fa strada il concetto di *augmented creativity* o di c.d. creatività aumentata. Vale a dire, di processi creativi in cui le capacità umane e quelle artificiali si integrano sinergicamente, generando risultati impossibili da conseguire per ciascun attore isolatamente. Questa prospettiva supera la dicotomia strumento-sostituto. In modo particolare,

l'IA non è più né mero strumento passivo né sostituto indipendente, ma *partner* collaborativo in processi creativi distribuiti.

Esempi embrionali di *augmented creativity* esistono già. Sistemi di IA che suggeriscono completamenti a musicisti che improvvisano; algoritmi che generano variazioni su temi proposti da pittori; chatbot che dialogano con scrittori aiutandoli a sviluppare trame e personaggi. In questi contesti, la creatività emerge dall'interazione iterativa: l'umano fornisce visione e direzione, l'IA genera proposte, l'umano seleziona e rielabora, l'IA raffina, e così via. Il risultato finale non è attribuibile semplicemente all'uno o all'altro ma al sistema collaborativo. Vien da sé che lo sviluppo della *augmented creativity* richiede ripensamento delle interfacce. Attualmente, l'interazione con l'IA avviene principalmente attraverso *prompt* testuali o parametri numerici, ovvero con modalità povere rispetto alla ricchezza dell'espressione umana. Tuttavia, le interfacce future potrebbero permettere di ricorrere ad una comunicazione multimodale – voce, gesto, espressione facciale, segnali cerebrali – in grado di catturare meglio intenzioni e sfumature. Sistemi di IA potrebbero apprendere stili e preferenze di collaboratori umani specifici, diventando *partner* sempre più sincronizzati e irrinunciabili.

Tuttavia, l'*augmented creativity* solleva anche moltissime preoccupazioni. Vi è rischio di dipendenza: creatori umani potrebbero atrofizzare le loro capacità native affidandosi eccessivamente all'assistenza algoritmica. Vi è, inoltre, il rischio di omogeneizzazione: se tutti utilizzassero stessi strumenti di IA, gli stili potrebbero convergere perdendo diversità. Affiora anche un pericolo alienazione: la separazione tra concezione e esecuzione, infatti, potrebbe impoverire l'esperienza creativa. Ovviamente, la gestione di questi rischi richiede consapevolezza critica e progettazione intenzionale di sistemi che amplificano – e non sostituiscono né impoveriscono – le capacità umane.

Un interrogativo finale, filosoficamente profondo, riguarda l'esistenza (o meno) di un nucleo irriducibilmente umano che resisterà all'automazione, anche nello scenario estremo di AGI. Questa domanda presenta un duplice risvolto: sia ontologico (che cosa costituisce l'umano?) sia normativo (che cosa dell'umano dovrebbe essere preservato?). Diverse sono le posizioni così come differenti sono le possibili risposte.

Una posizione essenzialista sostiene che esistono qualità specificamente umane – coscienza fenomenica, libero arbitrio, capacità di attribuire significato, mortalità – che nessun sistema artificiale potrà mai possedere perché tali prerogative sono legate alla natura biologica dell'essere umano. In questa prospettiva, per quanto sofisticata diventi l'IA, essa

rimarrà sempre fondamentalmente diversa e inferiore rispetto all'uomo. Invero, la creatività umana autentica, radicata in esperienza incarnata nel mondo, rimarrà insuperabile.

Una posizione funzionalista, al contrario, nega che vi sia qualcosa di speciale nella biologia umana. Ciò che conta, in realtà, sono le funzioni cognitive e comportamentali; se l'IA le replica, allora essa è intelligente e creativa quanto gli esseri umani. In questa prospettiva, la distinzione umano/artificiale perderà progressivamente significato mano a mano che le prestazioni convergeranno. L'unica cosa specificamente umana sarà la contingenza storica di essere stati i primi a sviluppare intelligenza su questo pianeta.

Una posizione intermedia, poi, riconosce che la questione ontologica (l'IA può essere identica all'intelligenza umana?) è meno importante della questione normativa (che valore attribuiamo alla specificità umana?). In modo particolare, anche ammettendo che l'IA possa replicare tutte le funzioni cognitive umane, ciò non implica che la distinzione perda valore morale e giuridico. Gli esseri umani meritano una considerazione speciale non perché possiedono capacità uniche in senso assoluto, ma in virtù della loro appartenenza alla famiglia degli esseri umani e della comunità morale e politica che integrano. Preservare spazio per la creatività umana, in questo senso, rappresenta un imperativo etico perché è funzionale al fiorire dell'umano e dell'umanità stessa.

In termini pratici, però, la tutela della dimensione dell'umano – e della sua peculiare creatività – richiede politiche che resistano alla mercificazione totale dell'ingegno. Ciò include: riserva di ambiti in cui solo la creazione umana è valorizzata e remunerata; educazione umanistica che sviluppa capacità non facilmente automatizzabili; sussidi e sostegno pubblico per arti e cultura non commerciali; regolamentazione che impedisce sostituzione completa di lavoro creativo umano. Queste politiche, ovviamente, non negano il valore della tecnologia, ma affermano che l'automazione deve servire l'umano fiorire, non già alla sua sostituzione.

Da qui, come in ogni momento di rivoluzione e di transizione, un rinnovato e quanto mai vigoroso appello al diritto (e al giurista) perché si ponga in ascolto dei cambiamenti, interpretandoli e dirigendoli al meglio secondo giustizia.



## CONCLUSIONI

Il percorso compiuto ha toccato vari snodi concettuali, tuttora oggetto di dibattito. Un primo snodo fondamentale riguarda la natura ontologica dell'IA generativa e il suo ruolo nell'ambito del processo creativo. L'analisi condotta ha evidenziato che l'IA generativa non può essere ridotta né alla categoria di mero strumento passivo, né a quella di autore pienamente autonomo, ma costituisce, piuttosto, un fenomeno genuinamente nuovo che richiede categorie concettuali specifiche. Essa è un *tertium genus*: vale a dire un'entità che esibisce forme di *agency* limitata, che produce *output* originali attraverso processi che incorporano elementi tanto deterministici quanto stocastici e che opera in spazi concettuali che possono trascendere la comprensione umana diretta. Questa constatazione ontologica ha conseguenze immediate per il diritto. Se l'IA non è né strumento né autore, le soluzioni normative non possono consistere semplicemente nell'applicazione estensiva delle categorie esistenti ma richiedono un'innovazione giuridica. Passando dalla *electronic personhood* alle forme di tutela *sui generis*, si è sottolineato che la soluzione più promettente consiste in un approccio graduato e pluralista, che attribuisca conseguenze giuridiche proporzionate al grado effettivo di autonomia e al contributo creativo dei diversi attori coinvolti nel processo generativo.

Un secondo tema nodale concerne la crisi del concetto di originalità quando applicato a produzioni algoritmiche. L'originalità, tradizionalmente intesa come espressione di scelte libere e creative dell'autore, incorpora presupposti antropologici che l'IA mette in discussione. Le opere generate dall'IA possono esibire novità statistica, complessità strutturale e valore estetico senza derivare da intenzionalità consapevole o da esperienza soggettiva incarnata. Una proposta interessante è quella di adottare il criterio del controllo creativo significativo. In pratica, l'originalità giuridicamente rilevante richiede che un agente umano identificabile abbia esercitato scelte creative specifiche e verificabili sul processo generativo – attraverso la determinazione del prompt, la selezione iterativa tra output alternativi, o la modificazione sostanziale del risultato – indipendentemente dal fatto che l'esecuzione tecnica sia stata affidata a un sistema algoritmico.

Un terzo argomento cruciale riguarda i fondamenti filosofici del diritto d'autore e la loro tenuta nell'era algoritmica. L'analisi svolta suggerisce che il diritto d'autore – se vuole sopravvivere nell'era dell'IA – necessita di una rifondazione teorica che integri elementi delle diverse tradizioni in un *framework* innovativo e coerente.

Un quarto snodo concettuale attiene la dimensione etica e distributiva. L'IA generativa, infatti, non pone solo problemi tecnico-giuridici ma solleva questioni fondamentali di giustizia: chi beneficia dell'innovazione tecnologica e chi ne sostiene i costi? Come prevenire che la concentrazione oligopolistica del potere tecnologico non si traduca in concentrazione del potere culturale e nella marginalizzazione di creativi umani? Come garantire che l'accesso agli strumenti creativi non riproduca o amplifichi disuguaglianze preesistenti? L'analisi ha mostrato che nessuna teoria della giustizia isolata – né quella rawlsiana, né quella seniana, né quella nozickiana – fornisca risposte univoche, ma che soltanto un approccio pluralista sensibile ai diversi valori in gioco risulta utile e percorribile.

Una quinta tematica riguarda le prospettive regolative. In modo particolare si è posto l'accento sui principi filosofico-giuridici trasversali – quello di precauzione, di sussidiarietà, di proporzionalità, di bilanciamento – che possono (e che devono) orientare la *governance* dell'IA.

Infine, si è volto lo sguardo alle traiettorie tecnologiche prevedibili e alle loro implicazioni per il diritto, per la filosofia e per l'etica. Scenari sui quali è necessario iniziare a riflettere affinché la società possa orientare consapevolmente lo sviluppo tecnologico piuttosto che subirlo passivamente.

Al di là dei risultati specifici concernenti il diritto d'autore, la ricerca ha messo in luce come l'intelligenza artificiale costituisca una sfida radicale per il diritto e soprattutto per la filosofia del diritto contemporanea. Una sfida che si gioca su molteplici livelli: ontologico, epistemologico, normativo e metodologico. Comprendere la natura e la portata di questa sfida è essenziale per valutare l'adeguatezza delle categorie giuridiche esistenti e per concorrere attivamente e proficuamente alla risemantizzazione di concetti e paradigmi.

Sul piano ontologico, l'IA mette in discussione quei binomi fondamentali sui quali il diritto si è edificato: persona/cosa, soggetto/oggetto, agente/strumento, naturale/artificiale. Il diritto moderno – pur avendo progressivamente esteso la personalità giuridica oltre i confini degli esseri umani individuali (società, enti pubblici, fondazioni) – ha mantenuto una distinzione netta tra soggetti capaci di azione giuridicamente rilevante e oggetti inerti. Oggi, l'IA – esibendo capacità di azione autonoma, apprendimento, adattamento e produzione di effetti non completamente

prevedibili – si colloca in una zona grigia che sfida tale distinzione. Non è né persona nel senso tradizionale, ma nemmeno cosa nel senso classico del termine. Questa ambiguità ontologica si riflette in difficoltà concettuali profonde. Se l'IA agisce, chi agisce effettivamente? Se produce danni, chi risponde? Se genera valore, a chi spetta?

Come è intuitivo a vacillare sono le categorie tradizionali di imputazione, di responsabilità, di titolarità. Categorie che presuppongono soggetti chiaramente identificabili dotati di volontà e intenzionalità. L'IA – operando attraverso processi algoritmici complessi che distribuiscono *agency* tra molteplici attori (programmatori, utilizzatori, sistemi stessi) – resiste a tale localizzazione. Inoltre, l'IA solleva interrogativi circa il bilanciamento tra diversi valori. Nell'ambito del diritto d'autore, ciò emerge chiaramente nella tensione tra innovazione tecnologica e tutela dei diritti degli autori, tra accesso universale alla conoscenza e incentivi alla creazione, tra libertà di ricerca e protezione della proprietà intellettuale.

Se così, vien da chiedersi quali prospettive si aprano per il futuro. Una prima direttrice – che potremmo definire di conservazione adattativa – prevede il mantenimento dell'architettura fondamentale del diritto d'autore attuale con aggiustamenti incrementali per incorporare l'IA. In pratica, si continuerebbero a riconoscere solo gli esseri umani come autori; opere generate con ausilio dell'IA riceverebbero protezione solo se l'intervento umano è sufficientemente significativo; opere generate autonomamente rimarrebbero nel dominio pubblico o riceverebbero tutela *sui generis* ridotta; eccezioni per text and data mining permetterebbero addestramento di modelli con salvaguardie appropriate. Questo scenario ha il vantaggio della continuità e della familiarità, minimizzando disruzioni al sistema consolidato. Tuttavia, rischia di risultare inadeguato se l'IA evolve verso forme sempre più autonome di creatività.

Una seconda direttrice, di trasformazione radicale, prevede ripensamento profondo dei fondamenti del diritto d'autore. Potrebbe emergere un diritto d'autore post-antropocentrico che riconosca forme di autorialità non-umana, valutando opere in base a caratteristiche intrinseche piuttosto che in base all'identità del creatore. Oppure potrebbe affermarsi un paradigma completamente diverso centrato sui diritti collettivi alla conoscenza piuttosto che sui diritti individuali degli autori, con la conoscenza e la cultura trattate come *commons* da gestire democraticamente. Questo quadro risponderebbe più radicalmente alle sfide dell'IA ma comporterebbe costi di transizione elevati e rischi di conseguenze impreviste.

Un terzo scenario, forse più realistico, prevede la coesistenza di regimi normativi diversi per ambiti diversi. Opere umane tradizionali conti-

nuerebbero a essere regolate dal *copyright* classico; creazioni assistite da IA riceverebbero forme di protezione intermedie calibrate sul grado di intervento umano; produzioni completamente automatizzate potrebbero essere soggette a regimi sui generis settoriali; ambiti specifici (ricerca scientifica, educazione, usi non commerciali) beneficerebbero di eccezioni ampie. Questa frammentazione, pur complicando il quadro normativo, permetterebbe flessibilità e sperimentazione.

Indipendentemente dallo scenario che si realizzerà, alcune direzioni appaiono auspicabili. In primo luogo, la regolamentazione dovrebbe essere *evidence-based*: basarsi cioè su dati empirici circa effetti dell'IA, piuttosto che su speculazioni ideologiche. In secondo luogo, la regolamentazione dovrebbe essere *adaptive*: capace di aggiustarsi rapidamente a sviluppi tecnologici imprevisi attraverso meccanismi di revisione periodica, sperimentazione controllata e *learning-by-doing*. In terzo luogo, la regolamentazione dovrebbe essere *participatory*: coinvolgere stakeholder diversi in processi deliberativi democratici piuttosto che affidare decisioni a tecnocrati o lasciare che il mercato decida. In quarto luogo, dovrebbe essere *principles-based*: fondata su principi normativi chiari (dignità umana, giustizia distributiva, accesso alla cultura) piuttosto che su regole tecniche dettagliate che rischiano l'obsolescenza. In quinto luogo, dovrebbe essere *internationally coordinated*: evitare frammentazione eccessiva attraverso armonizzazione progressiva basata su standard condivisi.

Realizzare questi obiettivi richiede notevoli trasformazioni non solo del diritto sostanziale ma anche delle istituzioni e dei processi attraverso cui il diritto viene elaborato, interpretato e applicato. Legislatori devono sviluppare capacità di comprendere tecnologie complesse e anticipare loro implicazioni. Giudici devono acquisire competenze tecniche sufficienti per valutare casi che coinvolgono IA. Autorità regolatorie devono dotarsi di expertise interdisciplinare e strumenti innovativi. Cittadini devono sviluppare *literacy* tecnologica che permetta partecipazione informata alle decisioni collettive. L'Università, come istituzione dedicata alla produzione e disseminazione di conoscenza, ha una responsabilità particolare in questo processo: deve formare professionisti capaci di navigare complessità dell'intersezione tra tecnologia, diritto e società; produrre ricerca che illumini scelte pubbliche; facilitare dialogo tra discipline, tra accademia e società, tra locale e globale.

In quest'ottica, il contributo che quest'opera intende fornire non è quello di prospettare specifiche soluzioni normative, ma di problematizzare filosoficamente le questioni sollevate dall'IA generativa. L'intelligenza artificiale non è semplicemente nuova tecnologia che richiede aggiu-

stamenti normativi tecnici. Essa pone interrogativi fondamentali circa la natura della creatività, il significato dell'autorialità, il valore dell'umano, la struttura della comunità politica. Affrontare tali interrogativi richiede risorse intellettuali che la tradizione filosofica mette a disposizione: capacità di analisi concettuale rigorosa, sensibilità per questioni normative profonde, apertura al dialogo interdisciplinare.

Se questa monografia avrà contribuito, anche modestamente, a rendere il dibattito pubblico sull'IA più consapevole filosoficamente, più rigoroso concettualmente, e più attento eticamente, essa avrà adempiuto alla propria funzione. L'intelligenza artificiale rappresenta probabilmente la sfida più significativa che l'umanità affronterà nel XXI secolo. Come risponderemo a tale sfida – se riusciremo a orientare la tecnologia verso il servizio dell'umano fiorire o se permetteremo che essa ci ridefinisca in modi che tradiscono i nostri valori più profondi – dipende dalle scelte che compiamo oggi. Tali scelte richiedono saggezza che solo l'integrazione di conoscenza tecnica, riflessione filosofica, e deliberazione democratica può fornire.



## BIBLIOGRAFIA

- AA.VV., *The Onlife Manifesto. Being Human in a Hyperconnected Era*, Cham, Springer, 2015.
- S. AARONSON, *Quantum Computing Since Democritus*, Cambridge, Cambridge University Press, 2013.
- R. ABBOTT, "I Think, Therefore I Invent: Creative Computers and the Future of Patent Law", in *Boston College Law Review*, 57(4), 2016, pp. 1079-1126.
- L.B. ALBERTI, *De pictura* (1435), in Id., *Opere volgari*, a cura di C. Grayson, Bari, Laterza, 1973.
- A.C. AMATO MANGIAMELI, *Natur@. Dimensioni della Biogiuridica*, Torino, Giappichelli, 2021.
- A.C. AMATO MANGIAMELI, "Sul diritto d'autore e la proprietà intellettuale, ovvero l'opera dell'ingegno tra dono, diritto e tecnica", in Id., *Informatica giuridica. Appunti e materiali ad uso di lezioni*, Torino, Giappichelli, 2015.
- A.C. AMATO MANGIAMELI, "Intelligenza artificiale generativa e Metaverso. A proposito di menti e corpi fusi con IA", in *Rivista di Filosofia del Diritto*, 1/2024, pp. 45-55.
- S. AMATO, "Darwin tra le macchine? L'intelligenza della tecnoscienza", in *Teoria e Critica della Regolazione Sociale*, 2/2022, pp. 13-30.
- S. AMATO, "Macchinazioni. Soggettività, agentività ed entità esotiche", in *L'Ircocervo*, 1/2024, pp. 289-306.
- H. ARENDT, *Vita activa. La condizione umana*, trad. it., Milano, Bompiani, 1964.
- ARISTOTELE, *Etica Nicomachea*, trad. it., Roma-Bari, Laterza, 2005.
- F. ARUTE et al., "Quantum Supremacy Using a Programmable Superconducting Processor", in *Nature*, 574, 2019, pp. 505-510.
- P. AUTERI, G. FLORIDIA, V. MANGINI, G. OLIVIERI, M. RICOLFI, P. SPADA, *Diritto industriale. Proprietà intellettuale e concorrenza*, Torino, Giappichelli, 2016.
- R. BALDWIN, M. CAVE, M. LODGE, *Understanding Regulation: Theory, Strategy, and Practice*, Oxford, Oxford University Press, 2012.
- R. BARTHES, "La mort de l'auteur", in *Manteia*, 5, 1967, pp. 12-17.
- J. BAUDRILLARD, *La società dei consumi*, trad. it., Bologna, il Mulino, 1974.
- Z. BAUMAN, *Modernità liquida*, Roma-Bari, Laterza, 2002.
- Z. BAUMAN, *Vite di scarto*, Roma-Bari, Laterza, 2004.
- S. BAYERN, "The Implications of Modern Business-Entity Law for the Regulation of Autonomous Systems", in *Stanford Technology Law Review*, 19, 2015, pp. 93-112.
- T.L. BEAUCHAMP, J.F. CHILDRESS, *Principles of Biomedical Ethics*, 8a ed., New York, Oxford University Press, 2019.
- W. BENJAMIN, *L'opera d'arte nell'epoca della sua riproducibilità tecnica*, trad. it., in Id., *Opere complete*, vol. VI, *Scritti 1934-1937*, trad. it., Torino, Einaudi, 2004, pp. 3-45.

- Y. BENKLER, *The Wealth of Networks: How Social Production Transforms Markets and Freedom*, New Haven, Yale University Press, 2006.
- J. BENTHAM, *An Introduction to the Principles of Morals and Legislation*, Oxford, Clarendon Press, 1789.
- L. BENTLY, B. SHERMAN, *Intellectual Property Law*, Oxford, Oxford University Press, 2014.
- M. BIAGIOLI, *Nature Engaged: Science in Practice from the Renaissance to the Present*, New York, Palgrave Macmillan, 2011.
- N. BOBBIO, *L'età dei diritti*, Torino, Einaudi, 1990.
- N. BOBBIO, *Teoria dell'ordinamento giuridico*, Torino, Giappichelli, 1960.
- M. BODEN, *The Creative Mind: Myths and Mechanisms*, London, Routledge, 2004.
- M. BOLDRIN, D.K. LEVINE, *Against Intellectual Monopoly*, Cambridge, Cambridge University Press, 2008.
- M. BOLDRIN, D.K. LEVINE, "The Case Against Patents", in *Journal of Economic Perspectives*, 27(1), 2013, pp. 3-22.
- G. BONGIOVANNI, G. SARTOR, C. VALENTINI, (a cura di), *Reasonableness and Law*, Dordrecht, Springer, 2009.
- J.L. BORGES, "La biblioteca de Babel" (1941), in ID., *Finzioni*, Torino, Einaudi, 1955.
- M. BORGHI, "Il diritto d'autore tra regime proprietario e interesse pubblico", in *Proprietà digitale: Diritti d'autore, nuove tecnologie e Digital Rights Management*, Milano, EGEA, 2006, pp. 1-21.
- D.F. BOUCHARD, (a cura di), *Language, Counter-Memory, Practice: Selected Essays and Interviews*, Ithaca (NY), Cornell University Press, 1977.
- J. BOYLE, "The Second Enclosure Movement and the Construction of the Public Domain", in *Law and Contemporary Problems*, 66(1/2), 2003, pp. 33-74.
- F. BRENTANO, *Psicologia dal punto di vista empirico*, trad. it., Bari, Laterza, 1997.
- F. BRENTANO, *Sull'origine della conoscenza morale*, trad. it., Brescia, Scholé, 2025.
- S. BREYER, "The Uneasy Case for Copyright: A Study of Copyright in Books, Photocopies, and Computer Programs", in *Harvard Law Review*, 84(2), 1970, pp. 281-351.
- M. BROOKS, *Quantum Computing: Progress and Prospects*, Washington DC, National Academies Press, 2019.
- M. BUSUIOC, D. CURTIN, M. ALMADA, "Reclaiming Transparency: Contesting the Logics of Secrecy within the AI Act", in *European Law Open*, 2(1), 2023, pp. 79-105.
- W.H. CAPITAN, D.D. MERRIL, (a cura di), *Art, Mind, and Religion*, Pittsburgh, University of Pittsburgh Press, 1967.
- N. CARLINI, F. TRAMÈR, E. WALLACE, M. JAGIELSKI, A. HERBERT-VOSS, K. LEE, A. ROBERTS, T.B. BROWN, D. SONG, Ú. ERLINGSSON, A. OPREA, C. RAFFEL, "Extracting Training Data from Large Language Models", in *Proceedings of the 30th USENIX Security Symposium*, USENIX Association, 2021, pp. 2633-2650.
- D.J. CHALMERS, *The Conscious Mind: In Search of a Fundamental Theory*, Oxford, Oxford University Press, 1996.

- A. CLARK, D.J. CHALMERS, "The Extended Mind", in *Analysis*, 58(1), 1998, pp. 7-19.
- R.H. COASE, "The Problem of Social Cost", in *Journal of Law and Economics*, 3, 1960, pp. 1-44.
- G.A. COHEN, *Self-Ownership, Freedom, and Equality*, Cambridge, Cambridge University Press, 1995.
- S.T. COLERIDGE, *Biographia Literaria*, Princeton, Princeton University Press, 1983.
- G. COMANDINI, "Arte, intelligenza artificiale e diritto d'autore: prime riflessioni", in *Rivista di diritto industriale*, 2023, pp. 3-28.
- CONGRESSIONAL RESEARCH SERVICE, "Artificial Intelligence: Overview, Recent Advances, and Considerations for the 118th Congress", R47644, 4 agosto 2023.
- D. COPE, *Virtual Music: Computer Synthesis of Musical Style*, Cambridge (MA), MIT Press, 2001.
- S. COTTA, *Diritto, persona, mondo umano*, Torino, Giappichelli, 1989.
- S. COTTA, *Il diritto nell'esistenza. Linee di ontofenomenologia giuridica*, Milano, Giuffrè, 1981.
- S. COTTA, *La sfida tecnologica*, Milano, Giuffrè, 1987.
- K. CRAWFORD, *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*, New Haven, Yale University Press, 2021.
- A. D'AGOSTINO, *Filosofia del diritto*, Torino, Giappichelli, 2006.
- F. D'AGOSTINO, *Bioetica nella prospettiva della filosofia del diritto*, Torino, Giappichelli, 1996.
- F. D'AGOSTINO, *Giustizia. Elementi per una teoria*, Cinisello Balsamo (MI), San Paolo, 2006.
- F. D'AGOSTINO, *Jus quia justum. Lezioni di filosofia del diritto e della religione*, Torino, Giappichelli, 2012.
- F. D'AGOSTINO, *Parole di giustizia*, Torino, Giappichelli, 2006.
- D.C. DENNETT, *The Intentional Stance*, Cambridge (MA), MIT Press, 1987.
- H.L. DREYFUS, S.E. Dreyfus, *Mind over Machine: The Power of Human Intuition and Expertise in the Era of the Computer*, New York, Free Press, 1986.
- H.L. DREYFUS, *What Computers Can't Do: A Critique of Artificial Reason*, New York, Harper & Row, 1972.
- H.L. DREYFUS, *What Computers Still Can't Do: A Critique of Artificial Reason*, Cambridge (MA), MIT Press, 1992.
- R. DWORKIN, *Law's Empire*, Cambridge (MA), Harvard University Press, 1986.
- R.A. EPSTEIN, "One Step Beyond Nozick's Minimal State: The Role of Forced Exchanges in Political Theory", in *Social Philosophy and Policy*, 22(1), 2005, pp. 286-313.
- M. FARINA, "Persona", in A.C. AMATO MANGIAMELI, G. SARACENI (a cura di), *Cento e una voce di Biogiuridica*, Torino, Giappichelli, 2024, pp. 347-351.
- M. FARINA, (a cura di), *Lo "Scudo Giuridico Adattivo". Intelligenza artificiale, cybersecurity e diritti fondamentali nell'ecosistema serics*, Torino, Giappichelli, 2026.
- M. FASAN, *Creatività e intelligenza artificiale*, Napoli, Edizioni Scientifiche Italiane, 2023.

- V. FELDMAN, "Does Learning Require Memorization? A Short Theory of a Long Tail", in *Proceedings of the 52nd Annual ACM Symposium on Theory of Computing (STOC)*, New York, ACM Press, 2020.
- L. FERRAJOLI, *Principia iuris. Teoria del diritto e della democrazia*, Roma-Bari, Laterza, 2007.
- M. FERRARIS, *Documentalità. Perché è necessario lasciar tracce*, Roma-Bari, Laterza, 2009.
- E. FESER, "Taxation, Forced Labor, and Theft", in *The Independent Review*, 5(2), 2000, pp. 219-235.
- M. FICINO, *De amore* (1469), in ID. *El libro dell'Amore*, Firenze, Olschki, 1987.
- L. FLORIDI, (a cura di), *The Onlife Manifesto. Being Human in a Hyperconnected Era*, Cham, Springer, 2015.
- L. FLORIDI, *La quarta rivoluzione. Come l'infosfera sta trasformando il mondo*, Milano, Cortina, 2017.
- L. FLORIDI, *The Ethics of Artificial Intelligence: Principles, Challenges and Opportunities*, Oxford, Oxford University Press, 2023.
- L. FLORIDI, *The Information: A Very Short Introduction*, Oxford, Oxford University Press, 2010.
- M. FOUCAULT, "Nietzsche, Genealogy, History", in D.F. BOUCHARD (a cura di), *Language, Counter-Memory, Practice: Selected Essays and Interviews*, Ithaca (NY), Cornell University Press, 1977, pp. 139-164.
- M. FOUCAULT, "Qu'est-ce qu'un auteur?", in *Bulletin de la Société Française de Philosophie*, 63(3), 1969, pp. 73-104.
- H. FRANKFURT, *Sulla verità*, trad. it., Milano, Feltrinelli, 2007.
- M. GAWDAT, *Superintelligenti. Come salvare il mondo dall'intelligenza artificiale*, trad. it., Milano, Rizzoli, 2022.
- O. VON GIERKE, *Das deutsche Genossenschaftsrecht*, vol. 1, Berlin, Weidmann, 1868.
- G. GOMETZ, "Intelligenza artificiale, profilazione e nuove forme di discriminazione", in *Teoria e Storia del Diritto Privato*, n. spec., 2022, pp. 1-38.
- K. GRACE, J. SALVATIER, A. DAFOE, B. ZHANG, O. EVANS, "Viewpoint: When Will AI Exceed Human Performance? Evidence from AI Experts", in *Journal of Artificial Intelligence Research*, 62, 2018, pp. 729-754.
- K. GRACE, J. SELBST, S. BEARD, O. SANDBERG, *AI Impacts Survey 2022: Expert Survey on Progress in AI*, Berkeley, Machine Intelligence Research Institute (MIRI), 2022.
- A. GUADAMUZ, "Do Androids Dream of Electric Copyright? Comparative Analysis of Originality in Artificial Intelligence Generated Works", in *Intellectual Property Quarterly*, 2, 2017, pp. 169-186.
- R. GUASTINI, *Interpretare e argomentare*, Milano, Giuffrè, 2011.
- J.P. GUILFORD, "Creativity", in *American Psychologist*, 5(9), 1950, pp. 444-454.
- J. HABERMAS, *Teoria dell'agire comunicativo*, trad. it., Bologna, il Mulino, 1986.
- J. HABERMAS, *Un nuovo mutamento strutturale della sfera pubblica e la politica deliberativa*, trad. it., Milano, Raffaello Cortina, 2023.

- J. HABERMAS, *Un nuovo mutamento strutturale della sfera pubblica e la politica deliberativa*, trad. it., Milano, Raffaello Cortina, 2023.
- M. HARDT, A. NEGRI, *Commonwealth*, Cambridge (MA), Harvard University Press, 2009.
- H.L.A. HART, *The Concept of Law*, Oxford, Oxford University Press, 1961.
- D. HARTMANN, J.R.L. DE PEREIRA, C. STREITBÖRGER, B. BERENDT, “Addressing the Regulatory Gap: Moving Towards an EU AI Audit Ecosystem Beyond the AI Act by Including Civil Society”, in *AI and Ethics*, 5, 2025, pp. 3617-3638.
- G.W.F. HEGEL, *Lineamenti di filosofia del diritto*, trad. it., Roma-Bari, Laterza, 1996.
- M.A. HELLER, R.S. EISENBERG, “Can Patents Deter Innovation? The Anticommons in Biomedical Research”, in *Science*, 280(5364), 1998, pp. 698-701.
- M.A. HELLER, “The Tragedy of the Anticommons: Property in the Transition from Marx to Markets”, in *Harvard Law Review*, 111(3), 1998, pp. 621-688.
- C. HESSE, “The Rise of Intellectual Property, 700 B.C.-A.D. 2000: An Idea in the Balance”, in *Daedalus*, 131(2), 2002, pp. 26-45.
- J. HUGHES, “The Philosophy of Intellectual Property”, in *Georgetown Law Journal*, 77(2), 1988, pp. 287-366.
- R.M. HURT, R.M. SCHUCHMAN, “The Economic Rationale of Copyright”, in *American Economic Review*, 56(1/2), 1966, pp. 421-432.
- E. HUSSERL, *Idee per una fenomenologia pura e per una filosofia fenomenologica*, trad. it., Torino, Einaudi, 2002.
- P. JASZI, “Toward a Theory of Copyright: The Metamorphoses of ‘Authorship’”, in *Duke Law Journal*, 1991(2), 1991, pp. 455-502.
- E. JIANG, K. NAOUS, A. RAMBACHAN, “Artificial Intelligence and Jobs: The Role of Demand”, *NBER Working Paper*, n. 31870, 2023.
- I. KANT, *Critica della capacità di giudizio*, trad. it., Milano, BUR, 2004.
- I. KANT, *Fondazione della metafisica dei costumi*, trad. it., Milano, Bompiani, 2003.
- H. KELSEN, *La dottrina pura del diritto*, trad. it., Torino, Einaudi, 1966.
- R. KILIAN, L. JÄCK, D. EBEL, “European AI Standards – Technical Standardisation and Implementation Challenges under the EU AI Act”, in *European Journal of Risk Regulation*, 16, 2025, pp. 1038-1062.
- P.W. KOH, P. LIANG, “Understanding Black-box Predictions via Influence Functions”, in *Proceedings of the 34th International Conference on Machine Learning (ICML)*, 2017.
- M. KOWALA, “Protection of Press Publishers in the Age of Generative AI – In Search of Legal Remedies to Adapt to the Pace of Technology”, in *IIC-International Review of Intellectual Property and Competition Law*, 55, 2024, pp. 1604-1623.
- R. KURZWEIL, *The Singularity Is Near: When Humans Transcend Biology*, New York, Viking, 2005.
- R. KURZWEIL, *The Singularity Is Nearer: When We Merge with AI*, New York, Viking, 2024.
- W.M. LANDES, R.A. POSNER, *The Economic Structure of Intellectual Property Law*, Cambridge (MA), Harvard University Press, 2003.

- D. LAZER et al., “The Science of Fake News”, in *Science*, 359(6380), 2018, pp. 1094-1096.
- J. LE DOUX, *Lunga storia di noi stessi. Come il cervello è diventato cosciente*, trad. it., Milano, Cortina, 2020.
- L. LESSIG, *Code and Other Laws of Cyberspace*, New York, Basic Books, 1999.
- L. LESSIG, *Free Culture: How Big Media Uses Technology and the Law to Lock Down Culture and Control Creativity*, New York, Penguin Press, 2004.
- L. LESSIG, “The Creative Commons”, in *Montana Law Review*, 65, 2004.
- J. LITMAN, *Digital Copyright*, Amherst, Prometheus Books, 2001.
- J. LOCKE, *Due trattati sul governo*, trad. it., Torino, UTET, 2010.
- J. LOVELOCK, B. APPELYARD, *Novacene. L'era dell'iperintelligenza*, trad. it., Torino, Bollati Boringhieri, 2020.
- N. LUHMANN, *Das Recht der Gesellschaft*, Frankfurt am Main, Suhrkamp, 1993.
- A. MACINTYRE, *After Virtue: A Study in Moral Theory*, Notre Dame (IN), University of Notre Dame Press, 2007.
- L.A. MAGGIO, K. STRANACK, “Understanding Creative Commons”, in *Academic Medicine*, 95(2), 2020, p. 322.
- G. MARCUS, E. DAVIS, *Rebooting AI: Building Artificial Intelligence We Can Trust*, New York, Pantheon Books, 2019.
- E. MORIN, *La Méthode 1: La Nature de la Nature*, Paris, Éditions du Seuil, 1977.
- E. MORIN, “Articuler les disciplines”, in *Carrefour des sciences. Actes du colloque du Comité national de la recherche scientifique «Interdisciplinarité»*, Paris, CNRS, 1990.
- M. MUSTONEN, “Copyleft-the Economics of Linux and other Open Source Software”, in *Information Economics and Policy*, 15(1), 2003, pp. 99-121.
- T. NAGEL, “What Is It Like to Be a Bat?”, in *The Philosophical Review*, 83(4), 1974, pp. 435-450.
- H. NISSENBAUM, “Accountability in a Computerized Society”, in *Science and Engineering Ethics*, 2(1), 1996, pp. 25-42.
- C. NOVELLI, F. CASOLARI, A. ROTOLO, M. TADDEO, L. FLORIDI, “Taking AI Risks Seriously: a New Assessment Model for the AI Act”, in *AI & Society*, 39, 2024, pp. 2493-2497.
- R. NOZICK, *Anarchy, State, and Utopia*, New York, Basic Books, 1974.
- R. NOZICK, *The Examined Life: Philosophical Meditations*, New York, Simon & Schuster, 1989.
- M.C. NUSSBAUM, *Creating Capabilities: The Human Development Approach*, Cambridge (MA), Harvard University Press, 2011.
- G. PASCUZZI, *Il diritto dell'era digitale*, Bologna, il Mulino, 2020.
- F. PASQUALE, *The Coded Gaze: Bias in Artificial Intelligence*, Cambridge (MA), MIT Press, 2021.
- PLATONE, *Ione*, in ID., *Tutti gli scritti*, Milano, Bompiani, 2000.
- I. PNG, “Copyright: A Plea for Empirical Research”, in *Review of Economic Research on Copyright Issues*, 3(2), 2006, pp. 3-13.

- G. PRIORA, “Il text and data mining nella Direttiva (UE) 2019/790: luci e ombre di una disciplina attesa”, in *Diritto dell'informazione e dell'informatica*, 2020, pp. 419-452.
- H. PUTNAM, “Psychological Predicates”, in W.H. CAPITAN, D.D. MERRIL (a cura di), *Art, Mind, and Religion*, Pittsburgh, University of Pittsburgh Press, 1967, pp. 37-48.
- J. RAWLS, *A Theory of Justice*, Cambridge (MA), Harvard University Press, 1971.
- J. RAWLS, *Political Liberalism*, New York, Columbia University Press, 1993.
- J. RAZ, *Between Authority and Interpretation: On the Theory of Law and Practical Reason*, Oxford, Oxford University Press, 2009.
- M. RICOLFI, M. VAN EECHOU, F. MORANDO, P. TZIAVOS, L. FERRAO, “The ‘Licensing’ of Public Sector Information”, in *Informatica e diritto*, 2011, pp. 129-146.
- M. RICOLFI, “Il diritto d'autore”, in N. ABRIANI, G. COTTINO, M. RICOLFI, *Diritto industriale*, Padova, CEDAM, 2001.
- M. RICOLFI, “License contracts, free software and creative commons”, in *Italian National Reports to the XIXth International Congress of Comparative Law*, Milano, Giuffrè, 2014, pp. 493-530.
- C.P. RIGAMONTI, “Deconstructing Moral Rights”, in *Harvard International Law Journal*, 47(2), 2006, pp. 353-412.
- S. RODOTÀ, *Il diritto di avere diritti*, Roma-Bari, Laterza, 2012.
- M. ROSE, *Authors and Owners: The Invention of Copyright*, Cambridge (MA), Harvard University Press, 1993.
- M.N. ROTHBARD, *For a New Liberty: The Libertarian Manifesto*, New York, Macmillan, 1973.
- P. SAMUELSON, “Generative AI Meets Copyright”, in *Science*, 381(6654), 2023, pp. 1293-1294.
- A. SANTOSUOSSO, *Diritto, scienza, nuove tecnologie*, Padova, CEDAM, 2016.
- G. SARTOR, *L'intelligenza artificiale e il diritto*, Torino, Giappichelli, 2022.
- F.C. VON SAVIGNY, *System des heutigen römischen Rechts*, Berlin, Veit, 1840.
- U. Scarpelli, *Cos'è il positivismo giuridico*, Napoli, Edizioni Scientifiche Italiane, 1997.
- F.W.J. SCHELLING, *Sistema dell'idealismo trascendentale*, trad. it., Milano, Bompiani, 1989.
- B. SCIALLA, “Le modifiche al Codice penale introdotte dalla normativa nazionale in materia di intelligenza artificiale (L. 132/2025)”, in *IRPA*, 25 marzo 2026.
- J.R. SEARLE, *Menti, cervelli e programmi. Un dibattito sull'intelligenza artificiale*, trad. it., Milano, Clup, 1984.
- J.R. SEARLE, *The Construction of Social Reality*, New York, Free Press, 1995.
- J.R. SEARLE, *The Rediscovery of the Mind*, Cambridge (MA), MIT Press, 1992.
- J.R. SEARLE, “Minds, Brains, and Programs”, in *Behavioral and Brain Sciences*, 3(3), 1980, pp. 417-424.
- A. SEN, *Commodities and Capabilities*, Amsterdam, North-Holland, 1985.

- A. SEN, *Development as Freedom*, New York, Knopf, 1999.
- A. SEN, *Inequality Reexamined*, Oxford, Clarendon Press, 1992.
- A. SEN, *The Idea of Justice*, Cambridge (MA), Harvard University Press, 2009.
- B. SENDHOFF et al. , (a cura di), *Creating Brain-Like Intelligence*, Berlin, Springer, 2009.
- T. SIMMONDS, “Common Knowledge? The Rise of Creative Commons Licensing”, in *Legal Information Management*, 10(3), 2010, pp. 162-165.
- A. SLOMAN, “Some Requirements for Human-Like Robots: Why the Recent Over-Emphasis on Embodiment Has Held Up Progress”, in B. SENDHOFF et al. (a cura di), *Creating Brain-Like Intelligence*, Berlin, Springer, 2009, pp. 248-277.
- L.B. SOLUM, “Legal Personhood for Artificial Intelligences”, in *North Carolina Law Review*, 70(4), 1992, pp. 1231-1287.
- C. SUNSTEIN, *#Republic: Divided Democracy in the Age of Social Media*, Princeton, Princeton University Press, 2017.
- C. SUNSTEIN, *The Ethics of Influence: Government in the Age of Behavioral Science*, Cambridge, Cambridge University Press, 2016.
- G. TARELLO, *L'interpretazione della legge*, Milano, Giuffrè, 1980.
- R.H. THALER, C.R. SUNSTEIN, *Nudge: Improving Decisions About Health, Wealth, and Happiness*, New Haven, Yale University Press, 2008.
- TOMMASO D'AQUINO, *La Somma Teologica*, vol. I, Bologna, Edizioni Studio Domenicano, 1988.
- R. TOWSE, (a cura di), *Copyright in the Cultural Industries*, Cheltenham, Edward Elgar, 2002.
- A.M. TURING, *Macchine calcolatrici e intelligenza*, a cura di D. Marconi, Torino, Einaudi, 2025.
- M. WIDDER, W. WEST, M. WHITTAKER, “Open (For Business): Big Tech, Concentrated Power, and the Political Economy of Open AI”, in *SSRN*, 2023.
- C.W. ZHU, “‘Copyleft’ Reconsidered: Why Software Licensing Jurisprudence Needs Insights from Relational Contract Theory”, in *Social & Legal Studies*, 22(3), 2013.

## FONTI NORMATIVE E GIURISPRUDENZIALI

- Andersen et al. v. Stability AI Ltd. et al.*, United States District Court for the Northern District of California, No. 3:23-cv-00201-WHO (W.H. Orrick, J.), azione depositata il 13 gennaio 2023, procedimento in corso al 1 gennaio 2026.
- Andy Warhol Foundation for the Visual Arts, Inc. v. Goldsmith*, 598 U.S. 508 (2023).
- Authors Guild v. Google, Inc.*, 804 F.3d 202 (2d Cir. 2015).
- Authors Guild, Inc. v. HathiTrust*, 755 F.3d 87 (2d Cir. 2014).

- Baker v. Selden*, United States Supreme Court, 101 U.S. 99 (1879).
- Feist Publications, Inc. v. Rural Telephone Service Co.*, United States Supreme Court, 499 U.S. 340 (1991).
- Huston c. Société Turner Entertainment*, Cour de Cassation (Francia), sentenza del 28 maggio 1991, in *Revue internationale du droit d'auteur*, n. 149, p. 197.
- Kelly v. Arriba Soft Corp.*, 336 F.3d 811 (9th Cir. 2003).
- Nichols v. Universal Pictures Corp.*, United States Court of Appeals, 45 F.2d 119 (1930).
- Perfect 10, Inc. v. Amazon.com, Inc.*, 508 F.3d 1146 (9th Cir. 2007).
- Thaler v. Perlmutter*, United States District Court for the District of Columbia, No. 1:22-cv-01564-BAH (B.A. Howell, J.), 687 F. Supp. 3d 140 (D.D.C. 2023), decisione del 18 agosto 2023.
- Thaler v. Secretary of State for Digital, Culture, Media and Sport*, High Court of Justice (England and Wales), EWHC 2412 (Admin), 21 September 2023.
- European Parliament, Resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)), P8\_TA(2017)0051.
- Copyright, Designs and Patents Act 1988, c. 48, s. 9(3).
- Direttiva (UE) 2019/790 del Parlamento europeo e del Consiglio, 17 aprile 2019, sul diritto d'autore e sui diritti connessi nel mercato unico digitale.
- Executive Order 14110 of October 30, 2023, *Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence*.
- Executive Order 14148 of January 20, 2025, *Initial Rescissions of Harmful Executive Orders and Actions*, in *Federal Register*, 90 Fed. Reg. 8233 (28 gennaio 2025).
- Commissione europea, *An agile Digital Rulebook for the EU*, in *Shaping Europe's digital future*.
- Copyright Act of Japan, art. 30-4 (testo consolidato nella traduzione ufficiale del Japanese Law Translation Database System).
- Cyberspace Administration of China et al., *Interim Measures for the Management of Generative Artificial Intelligence Services*, 10 luglio 2023.
- Dichiarazione di Rio sull'ambiente e lo sviluppo, Rio de Janeiro: ONU, 1992.
- IMDA – AI Verify Foundation, *Proposed Model AI Governance Framework for Generative AI*, 16 gennaio 2024.
- Infopaq International A/S v. Danske Dagblades Forening*, CGUE, sentenza del 16 luglio 2009, causa C-5/08, EU:C:2009:465.
- ISO/IEC TR 24028:2020 – Information technology – Artificial intelligence – Overview of trustworthiness in artificial intelligence, Ginevra, International Organization for Standardization, 2020.
- ISO/IEC 38507:2022 – Information technology – Governance of IT – Governance implications of the use of artificial intelligence by organizations, Ginevra, International Organization for Standardization, 2022.
- ISO/IEC 23894:2023 – Information technology – Artificial intelligence – Guidance on risk management, Ginevra, International Organization for Standardization, 2023.

- OECD, *Recommendation of the Council on Artificial Intelligence*, OECD/LEGAL/0449, 22 maggio 2019, agg. 2023.
- Painer, Eva-Maria v. Standard Verlags GmbH and Others*, CGUE, sentenza del 1 dicembre 2011, causa C-145/10, EU:C:2011:798.
- Republic of Korea, Framework Act on the Development of Artificial Intelligence and the Creation of a Foundation for Trust, Act No. 20676, 21 gennaio 2025, effective 22 gennaio 2026.
- Legge 23 settembre 2025, n. 132, Disposizioni in materia di intelligenza artificiale.
- NITI Aayog, Responsible AI for All: Strategy for India, 2021.
- Regolamento (UE) 2024/1689 del Parlamento europeo e del Consiglio, 13 giugno 2024, recante regole armonizzate sull'intelligenza artificiale (AI Act).
- UNESCO, *Recommendation on the Ethics of Artificial Intelligence*, Paris, UNESCO, 2021.
- European Group on Ethics in Science and New Technologies, *Statement on Artificial Intelligence, Robotics and "Autonomous" Systems*, Brussels, European Commission, 2018.

## INDICE ANALITICO

- Aaronson, S. *Cap. VIII*  
Accountability *Cap. V, VII, C.*  
Actor-Network Theory (Latour) *Cap. I*  
AGI (Intelligenza artificiale generale)  
*Cap. VIII, C.*  
Agency (dell'IA) *Cap. III, IV, C.*  
AI Act (Reg. UE 2024/1689) *Cap. VII, C.*  
AI Impacts Survey *Cap. VIII*  
Alberti, L.B. *Cap. I*  
Algoritmo *Cap. III, IV, V*  
Amato Mangiameli, A.C. *Cap. II, V, VIII*  
Analogia fotocamera/sintetizzatore  
*Cap. III*  
Androcentrismo *v. Antropocentrismo*  
Antinomie del diritto d'autore *Cap. II, C.*  
Antropocentrismo *Cap. II, IV, VIII, C.*  
Approccio pluralista *Cap. I, V, VI, C.*  
Arendt, H. *Cap. I, Int.*  
Aristotele *Cap. I, V*  
Arute, F. et al. *Cap. VIII*  
Augmented creativity *Cap. VIII*  
Autenticità dell'opera *Cap. V, VIII*  
Autore *Cap. II, III, IV, C.*  
Autorialità *Cap. II, III, IV, C.*  
Autonomia (principio di) *Cap. V*  
Baldwin, R. - Cave, M. - Lodge, M.  
*Cap. VII*  
Barthes, R. *Cap. I*  
Beauchamp, T.L. - Childress, J.F. *Cap. V*  
Beneficenza (principio di) *Cap. V*  
Benjamin, W. *Cap. V, Int.*  
Bentham, J. *Cap. II, V*  
Bias algoritmico *Cap. V, VI*  
Bilanciamento *Cap. IV, VII, C.*  
Black box *Cap. V, VII*  
Boden, M. *Cap. I, III*  
Boldrin, M. - Levine, D.K. *Cap. VI, VIII*  
Borghi, M. *Cap. VIII*  
Brain-Computer Interfaces (BCI)  
*Cap. VIII*  
Brentano, F. *Cap. III*  
Brooks, M. *Cap. VIII*  
Busuioc, M. - Curtin, D. - Almada, M.  
*Cap. VII*  
Capabilities (approccio delle) *Cap. VI*  
Carlini, N. et al. *Cap. V*  
Causa formale (aristotelica) *Cap. I*  
Chalmers, D.J. *Cap. I, III*  
ChatGPT *Cap. III, IV, Int.*  
Clausola lockiana della sufficienza  
*Cap. II, VI*  
Co-autorialità *Cap. IV, C.*  
Cohen, G.A. *Cap. VI*  
Coleridge, S.T. *Cap. I*  
Commons / Bene comune *Cap. VIII, C.*  
Computer-generated works (CDPA  
1988) *Cap. IV, VII*  
Consequenzialismo *Cap. V*  
Controllo creativo significativo *Cap. III, IV, C.*  
Convenzione di Berna *Cap. II*  
Cope, D. *Cap. I*  
Copyleft *Cap. VIII*  
Copyright Act of Japan (art. 30-4)  
*Cap. VII*  
Crawford, K. *Cap. VI*

- Creatività *Cap. I, II, III, VIII, Int., C.*  
 Creatività algoritmica *Cap. I, III*  
 Creatività aumentata *v. Augmented creativity*  
 Creatività combinatoria / esplorativa / trasformazionale *Cap. I*  
 Creative Commons *Cap. VIII*  
 Creatio ex nihilo *Cap. I*  
 D'Agostino, F. *Cap. I, V*  
 DALL-E *Cap. III, Int.*  
 Dataset di addestramento *Cap. III, V, VII*  
 Deepfakes *Cap. V, VII*  
 Dennett, D.C. *Cap. III*  
 Deontologia kantiana *Cap. V*  
 Dichiarazione di Rio (1992) *Cap. VII*  
 Dignità umana *Cap. I, IV, V, VIII, C.*  
 Diritti fondamentali *Cap. IV, VII*  
 Diritti morali *Cap. II, IV*  
 Diritto alla cultura *Cap. IV, C.*  
 Diritto d'autore *Cap. II, VII, VIII, C.*  
 Diritto d'autore post-antropocentrico *Cap. VIII, C.*  
 Documentalità (teoria della) *Cap. I, II, III*  
 Dominio pubblico *Cap. II, VII, VIII, C.*  
 Dreyfus, H.L. *Cap. III*  
 Droit d'auteur *Cap. II*  
 Electronic personhood *Cap. IV, C.*  
 EMI (Experiments in Musical Intelligence) *Cap. I*  
 Entanglement quantistico *Cap. VIII*  
 Epstein, R.A. *Cap. VI*  
 Equilibrio riflessivo *v. Reflective equilibrium*  
 Etica della virtù *Cap. V*  
 Etica dell'IA *Cap. V*  
 Extended Mind (teoria della) *Cap. I*  
 Fair use *Cap. VII*  
 Fasan, M. *Cap. III, V*  
 Feser, E. *Cap. VI*  
 Ficino, M. *Cap. I*  
 Fissazione dell'output *Cap. III*  
 Floridi, L. *Cap. I, III, V*  
 Fodor, J. *Cap. III*  
 Foucault, M. *Cap. I, II*  
 Funzionalismo *Cap. III, IV*  
 Gawdat, M. *Cap. VIII*  
 Genio romantico *Cap. I, II*  
 Giustizia distributiva *Cap. VI, Int., C.*  
 Grace, K. et al. *Cap. VIII*  
 Guadamuz, A. *Cap. III*  
 Guilford, J.P. *Cap. I*  
 Habermas, J. *Cap. III, V, VI*  
 Hardt, M. - Negri, A. *Cap. VIII*  
 Hegel, G.W.F. *Cap. II, V*  
 Heidegger, M. *Int.*  
 Heller, M.A. *Cap. II*  
 Hughes, J. *Cap. II, VI*  
 Husserl, E. *Cap. III*  
 Idea/espressione (distinzione) *Cap. II, III*  
 Imperativo categorico *Cap. V*  
 Infopaq (sentenza) *Cap. I, II*  
 Infosfera *Cap. V*  
 Intelligenza artificiale generale (AGI) *v. AGI*  
 Intelligenza artificiale quantistica *Cap. VIII*  
 Intenzionalità *Cap. III, IV*  
 Intenzionalità mediata *Cap. III, IV*  
 Interdisciplinarietà *Cap. I, V, VIII, Int.*  
 Interfacce cervello-computer (BCI) *v. Brain-Computer Interfaces*  
 Jonas, H. *Int.*  
 Jaszi, P. *Cap. I*  
 Kant, I. *Cap. I, V*  
 Kelsen, H. *Cap. I*  
 Kilian, R. - Jäck, L. - Ebel, D. *Cap. VII*  
 Koh, P.W. - Liang, P. *Cap. VI*  
 Kowala, M. *Cap. VII*  
 Kurzweil, R. *Cap. VIII*  
 Latour, B. *Cap. I*  
 Lazer, D. et al. *Cap. V*  
 Le Doux, J. *Cap. VIII*  
 Lessig, L. *Cap. II, VII, VIII*  
 Libertà di espressione *Cap. IV*  
 Libertà processuale (Sen) *Cap. VI*

- Licensing collettivo esteso (ECL) *Cap. VI*
- Litman, J. *Cap. II*
- Locke, John *Cap. II, V*
- Lovelock, J. - Appleyard, B. *Cap. VIII*
- Luhmann, N. *Cap. I*
- Machine learning *Cap. III*
- MacIntyre, A. *Cap. V*
- Marcus, G. - Davis, E. *Cap. VIII*
- Memorizzazione dei dati di training  
*Cap. V*
- Mercificazione della creatività *Cap. V*
- Midjourney *Cap. III, Int.*
- Modelli generativi *Cap. III*
- Morin, E. *Cap. I*
- Mustonen, M. *Cap. VIII*
- Nagel, T. *Cap. III*
- Neuralink *Cap. VIII*
- Nissenbaum, H. *Cap. V*
- Non maleficenza (principio di) *Cap. V*
- Novelli, C. et al. *Cap. VII*
- Nozick, Robert *Cap. II, VI*
- Nussbaum, M.C. *Cap. V*
- Oligopolio tecnologico *Cap. VI, Int., C.*
- Onlife *Cap. V*
- Ontologia sociale *Cap. I*
- Opacità algoritmica *Cap. III, V, C.*
- Open source *Cap. VI, VII*
- Opt-out (riserva dei diritti) *Cap. IV, VII*
- Originalità *Cap. II, III, C.*
- Painer (sentenza) *Cap. I, II*
- Personalità giuridica *Cap. IV, C.*
- Phronesis (prudenza) *Cap. V*
- Platone *Cap. I*
- Poiesis / Praxis *Cap. I*
- Post-antropocentrismo *v. Diritto d'autore post-antropocentrico*
- Precauzione (principio di) *Cap. VII*
- Principialismo etico *Cap. V*
- Progetto HOLMES *Int., Cap. I-VIII, C.*
- Prompt *Cap. III, IV*
- Proporzionalità *Cap. VII*
- Proprietà intellettuale *Cap. II, VI, VIII*
- Putnam, H. *Cap. I, III*
- Qualia *Cap. III*
- Qubit *Cap. VIII*
- Rawls, John *Cap. VI*
- Reddito di base universale (UBI) *Cap. VI*
- Reductio della creatività a processo  
computazionale *Cap. I, III*
- Reflective equilibrium *Cap. V*
- Regulatory sandboxes *Cap. VII*
- Remunerazione degli autori *Cap. VI*
- Responsabilità distribuita *Cap. V, C.*
- Responsabilità etica *Cap. V*
- Ricolfi, M. *Cap. VIII*
- Rose, M. *Cap. I*
- Rothbard, M.N. *Cap. VI*
- Samuelson, P. *Cap. III, VII*
- Santosuosso, A. *Cap. V*
- Sartor, G. *Cap. V*
- Scarpelli, U. *Cap. I*
- Schelling, F.W.J. *Cap. I*
- Scialla, B. *Cap. VII*
- Searle, J.R. *Cap. I, III*
- Sen, Amartya *Cap. VI*
- Serialità creativa *Cap. III*
- Singolarità tecnologica *Cap. VIII*
- Sloman, A. *Cap. I*
- Soft law *Cap. V, VII*
- Soggettività giuridica *Cap. IV, C.*
- Solum, L.B. *Cap. II, IV*
- Spazio latente (modelli generativi)  
*Cap. III*
- Stable Diffusion *Cap. VI, Int.*
- Standardizzazione tecnica *Cap. VII*
- Stanza cinese (esperimento mentale)  
*Cap. I, III*
- Statuto di Anna (1710) *Cap. II*
- Status (fatti istituzionali) *Cap. I*
- Sussidiarietà *Cap. VII*
- Sunstein, C. *Cap. VI*
- Tarello, G. *Cap. I*
- Tertium genus *Cap. III, C.*
- Text and data mining (TDM) *Cap. II, IV, VII*
- Thaler v. Perlmutter *Cap. IV*

- Thaler v. Secretary of State (UK) *Cap.* IV, VII
- Teoria della documentalità *v.* *Documentalità*
- Teoria romantica dell'autore *Cap.* I, II, IV
- Tommaso d'Aquino *Cap.* I
- Training set *v.* *Dataset di addestramento*
- Trasparenza *Cap.* V, VII
- Trasformative use *Cap.* VII
- Turing, A.M. *Cap.* I
- Tutela sui generis *Cap.* IV, VII, C.
- UBI *v.* *Reddito di base universale*
- Urheberrecht *Cap.* II
- Utilitarismo *Cap.* II, V
- Velo di ignoranza (Rawls) *Cap.* VI
- Watermarking *Cap.* V
- Weber, M. *Int.*
- Widder, M. - West, W. - Whittaker, M. *Cap.* VI, VII
- Zhu, C.W. *Cap.* VIII
- Zombie filosofico (Chalmers) *Cap.* III

## GLOSSARIO DEI TERMINI TECNICI

**Accountability:** Responsabilità nel rendere conto delle proprie azioni o decisioni, con possibilità di essere chiamati a rispondere davanti a un'autorità o al pubblico. Nel contesto dell'IA generativa implica trasparenza sul funzionamento dei sistemi, meccanismi di controllo ex ante ed ex post e rimedi effettivi per chi subisce danni.

**Actor-Network Theory (ANT):** Approccio teorico elaborato da Bruno Latour che considera la creatività non come atto isolato di un soggetto sovrano, bensì come emergenza distribuita all'interno di una rete composta di attori umani e non-umani. Applicata all'IA, suggerisce di riconsiderare la dicotomia creatore/strumento.

**Agency:** Capacità di un soggetto (umano o artificiale) di agire autonomamente, prendere decisioni e produrre effetti nel mondo. Per l'IA generativa si distingue tra agency piena (intenzionalità, autonomia, razionalità) e agency debole (capacità funzionale senza coscienza), con rilevanti conseguenze per l'attribuzione di responsabilità e diritti.

**AGI (Intelligenza Artificiale Generale):** v. Intelligenza Artificiale Generale.

**AI Act:** Regolamento (UE) 2024/1689 del 13 giugno 2024 che stabilisce regole armonizzate per l'intelligenza artificiale nell'Unione Europea, primo quadro normativo orizzontale e vincolante sull'IA a livello mondiale. Basato su un approccio risk-based, prevede obblighi specifici per i modelli di uso generale (GPAI) e introduce meccanismi di accountability, trasparenza e regulatory sandboxes.

**Algoritmo:** Sequenza finita di istruzioni o regole ben definite che permettono di risolvere un problema o compiere un'operazione. Nel contesto dell'IA, insieme di procedure computazionali che elaborano dati per produrre output. La complessità degli algoritmi di deep learning rende spesso opaco il processo decisionale interno.

**Aura (Benjamin):** Concetto elaborato da Walter Benjamin per designare l'unicità irripetibile dell'opera d'arte originale, legata al suo hic et nunc. La riproducibilità tecnica dissolve l'aura; l'IA generativa radicalizza questo fenomeno producendo opere a costo marginale tendente a zero e rendendo obsoleta la distinzione tra originale e copia.

**Augmented creativity (Creatività aumentata):** Processo creativo in cui capacità umane e artificiali si integrano sinergicamente, generando risultati impos-

sibili per ciascun attore isolatamente. Supera la dicotomia strumento/sostituto: l'IA non è né strumento passivo né autore indipendente, ma partner collaborativo in processi creativi distribuiti.

**Bias algoritmico:** Distorsione sistematica nei risultati prodotti da un sistema di IA, derivante da dati di addestramento non rappresentativi o da scelte progettuali che incorporano pregiudizi culturali, di genere o economici. Solleva questioni di giustizia distributiva e non discriminazione.

**Black box:** v. Opacità algoritmica.

**Brain-Computer Interface (BCI):** Interfaccia cervello-computer: sistema che permette comunicazione diretta tra cervello biologico e dispositivi elettronici, bypassando canali sensoriali e motori tradizionali. Le BCI avanzate potrebbero consentire ibridazione tra intelligenza biologica e artificiale, sollevando questioni radicali sull'identità personale e l'attribuzione dei diritti d'autore.

**Capabilities approach (Approccio delle capacità):** Teoria della giustizia elaborata da Amartya Sen che valuta il benessere in termini di libertà sostanziali (capabilities) effettivamente disponibili a ciascuna persona, non di beni primari o utilità aggregate. Applicato all'IA, sposta il criterio dalla disponibilità formale degli strumenti alla capacità effettiva di esercitare creatività autonoma.

**ChatGPT:** Sistema di IA generativa per il linguaggio naturale sviluppato da OpenAI, basato sull'architettura transformer GPT. Simbolo della svolta nell'IA generativa accessibile al grande pubblico (2022), ha dimostrato capacità linguistiche che sfidano i confini tradizionali tra intelligenza umana e artificiale.

**Clausola lockiana della sufficienza:** Condizione elaborata da John Locke secondo cui l'appropriazione privata di risorse è legittima solo se lascia agli altri risorse sufficienti e di qualità altrettanto buona. Applicata alla proprietà intellettuale, solleva interrogativi sulla compatibilità dell'addestramento massivo su opere altrui con tale condizione.

**Commons / Bene comune:** Risorse o conoscenze prodotte socialmente che non possono essere privatizzate senza ingiustizia (Hardt e Negri). Nel contesto dell'IA, la proposta di trattare opere culturali e conoscenza come commons da gestire democraticamente si contrappone ai modelli di proprietà esclusiva.

**Computer-generated works:** Categoria prevista dalla section 9(3) del Copyright, Designs and Patents Act 1988 (UK) che attribuisce la paternità dell'opera alla persona che ha predisposto le condizioni necessarie per la creazione (arrangements necessary for the creation of the work), consentendo protezione anche in assenza di autore umano identificabile.

**Controllo creativo significativo:** Criterio proposto per determinare l'autorialità nell'era dell'IA: l'utilizzatore acquisisce la titolarità dei diritti nella misura in cui

esercita scelte creative specifiche e verificabili sul processo generativo (specificità dei prompt, iterazione degli output, selezione consapevole, modificazione del risultato).

**Copyleft:** Modello di licenza che, facendo leva sul diritto d'autore, consente uso, modifica e redistribuzione dell'opera a condizione che le opere derivate siano diffuse con termini analoghi, impedendo appropriazioni esclusive di sviluppi successivi.

**Copyright:** Termine di common law equivalente al diritto d'autore, che designa l'insieme di diritti esclusivi riconosciuti al creatore di un'opera dell'ingegno. A differenza del *droit d'auteur* continentale, il copyright è tradizionalmente orientato agli aspetti economici più che alla dimensione personalistica.

**Creative Commons:** Sistema di licenze pubbliche standardizzate che consente all'autore di modulare anticipatamente le condizioni di utilizzazione dell'opera, favorendone la circolazione e il riuso entro limiti predeterminati. Alternativa al modello proprietario tradizionale particolarmente rilevante nell'era dell'IA.

**DALL-E:** Sistema di IA generativa per immagini sviluppato da OpenAI, capace di produrre immagini di alta qualità a partire da descrizioni testuali. Tra i sistemi che hanno segnato la svolta nell'IA generativa accessibile al pubblico nel 2022.

**Dataset di addestramento (Training set):** Insieme di dati (testi, immagini, audio) utilizzato per addestrare un modello di IA generativa. Può contenere miliardi di esempi, spesso protetti da copyright. La liceità dell'uso di opere protette per l'addestramento è una delle questioni giuridiche più controverse nel diritto d'autore contemporaneo.

**Deep learning:** Sottocampo del machine learning basato su reti neurali artificiali con molteplici livelli (strati nascosti) che apprendono rappresentazioni gerarchiche dei dati. Tecnologia alla base dei principali sistemi di IA generativa.

**Deepfake:** Contenuto audiovisivo generato o manipolato attraverso IA in modo da risultare autentico pur essendo falso. Rappresenta un attacco all'infrastruttura epistemica condivisa, violando le condizioni di possibilità dell'agire comunicativo (Habermas). Solleva questioni di diritto all'immagine, reputazione, privacy e diritto d'autore.

**Dignità umana:** Valore intrinseco e inviolabile di ogni essere umano, fondamento dei diritti umani fondamentali. Principio cardine degli ordinamenti costituzionali moderni (art. 1 Carta dei diritti fondamentali UE). Opera come limite invalicabile all'estensione della soggettività giuridica all'IA: le macchine possono ricevere riconoscimento normativo funzionale ma non equivalenza morale con le persone fisiche.

**Diritti morali:** Diritti non patrimoniali dell'autore che tutelano il legame personale tra creatore e opera: diritto alla paternità, all'integrità, alla divulgazione e al pentimento. Fondati sulla teoria hegeliana della personalità, presuppon-

gono un legame individuale indissolubile tra autore umano e opera, messo in crisi dalla creatività algoritmica.

**Documentalità (teoria della):** Teoria elaborata da Maurizio Ferraris secondo cui la formula ‘oggetto sociale = atto iscritto’ descrive la struttura ontologica degli oggetti giuridicamente rilevanti. Applicata all’IA, implica che il diritto d’autore sorge con la fissazione dell’output su supporto e non in momenti anteriori (addestramento, formulazione del prompt).

**Dominio pubblico:** Insieme delle opere non più protette da diritto d’autore (per scadenza del termine) o mai proteggibili (idee, fatti, opere prive di originalità). Spazio essenziale per la creatività derivativa e l’accesso universale alla cultura. L’IA generativa solleva il rischio di privatizzazione del dominio pubblico attraverso la protezione massiva di output algoritmici.

**Droit d’auteur:** Sistema francese del diritto d’autore, espressione della tradizione giuridica continentale che, a differenza del copyright anglosassone, enfatizza il legame personale tra autore e opera e riconosce diritti morali inalienabili fondati sulla teoria hegeliana della personalità.

**Electronic personhood:** Proposta di attribuire personalità giuridica a sistemi artificiali sufficientemente sofisticati, analogamente a quanto avviene per le persone giuridiche. Esplorata dal Parlamento europeo nella risoluzione del 16 febbraio 2017 sulla robotica. Solleva obiezioni legate alla mancanza di interessi propri, all’impossibilità di esercitare diritti morali e al rischio di svilire la dignità umana.

**Enthousiasmós:** Ispirazione divina nella concezione greca antica della creatività: il poeta è strumento passivo delle Muse, non creatore autonomo. Concezione teurgica che ha dominato la cultura occidentale fino al Rinascimento, poi progressivamente sostituita dalla visione antropocentrica del genio individuale.

**Etica della virtù:** Approccio etico di matrice aristotelica che pone al centro la formazione del carattere morale degli agenti e lo sviluppo di disposizioni virtuose. Applicata all’IA, richiede che progettisti, sviluppatori e utilizzatori coltivino virtù come prudenza (phronesis), giustizia, temperanza e onestà intellettuale, piuttosto che limitarsi all’applicazione di regole esterne.

**Fair use:** Dottrina del diritto statunitense (17 U.S.C. §107) che permette l’utilizzo di opere protette senza autorizzazione sulla base di una valutazione case by case di quattro fattori: scopo e carattere dell’uso, natura dell’opera, quantità utilizzata, effetto sul mercato. Il carattere trasformativo (transformative use) è elemento chiave, come ridimensionato dalla Corte Suprema in *Andy Warhol Foundation v. Goldsmith* (2023).

**Fatti bruti / fatti istituzionali:** Distinzione elaborata da John Searle nell’ontologia sociale: i fatti bruti esistono indipendentemente dall’osservatore; i fatti

istituzionali sono costituiti da regole condivise. Il diritto d'autore non protegge proprietà metafisiche immanenti (fatti bruti) ma costituisce lo status giuridico dell'opera (fatto istituzionale) mediante regole costitutive.

**Fine-tuning:** Processo di ulteriore addestramento di un modello pre-addestrato su dati specifici per adattarlo a un compito o dominio particolare, migliorandone le prestazioni con minore investimento computazionale rispetto all'addestramento da zero.

**Fissazione dell'output:** Momento in cui l'output di un sistema di IA generativa viene registrato su un supporto materiale stabile. Secondo la teoria della documentalità (Ferraris), costituisce il momento costitutivo rilevante per il sorgere del diritto d'autore: prima della fissazione l'output è fatto brutto potenziale; con la fissazione diventa oggetto sociale su cui il diritto può operare.

**Foundation model:** Modello di IA di grandi dimensioni addestrato su dataset vasti e diversificati, che può essere adattato (fine-tuned) per molteplici compiti specifici. GPT, BERT, LLaMA sono esempi di foundation models che costituiscono l'infrastruttura tecnologica dell'IA generativa contemporanea.

**Funzionalismo:** Posizione filosofica secondo cui gli stati mentali sono definiti non dalla loro costituzione materiale ma dal ruolo causale che svolgono nel sistema cognitivo complessivo. Applicato all'IA, implica che se un sistema implementa gli stessi ruoli funzionali di stati mentali umani, possiede intenzionalità genuina indipendentemente dal substrato fisico.

**Generative AI (IA generativa):** Sistemi di intelligenza artificiale capaci di generare autonomamente nuovi contenuti (testi, immagini, audio, video, codice) che presentano caratteristiche statisticamente coerenti con i dati di addestramento ma non sono copie esatte di alcun elemento del corpus. Si distinguono dall'IA tradizionale per la capacità di produrre output originali attraverso interpolazione nello spazio latente delle rappresentazioni apprese.

**GPAI (General Purpose AI model):** Modello di IA per finalità generali: sistema addestrato su vasti dati che può svolgere un'ampia gamma di compiti diversi. Categoria introdotta dall'AI Act (Reg. UE 2024/1689, Chapter V) con obblighi specifici di trasparenza, documentazione tecnica e rispetto del copyright, con ulteriori obblighi per i modelli con rischio sistemico.

**H-creativity / P-creativity:** Distinzione elaborata da Margaret Boden tra creatività psicologica (P-creativity: idee nuove per il soggetto) e creatività storica (H-creativity: idee nuove per l'intera umanità o un campo culturale). Solo la H-creativity ha rilevanza culturale e, di riflesso, giuridica.

**Idea/espressione (distinzione):** Principio fondamentale del diritto d'autore per cui la protezione copre le forme espressive attraverso cui le idee sono manifestate,

ma non le idee stesse. Fondamento del libero flusso delle informazioni e della creatività derivativa. L'addestramento su opere protette pone la questione se l'IA estragga solo idee non proteggibili o si appropri anche di strutture espressive tutelate.

**Infosfera:** Concetto elaborato da Luciano Floridi per designare l'ambiente complessivo costituito da tutti gli agenti informazionali e le loro interazioni, nel quale gli esseri umani vivono sempre più immersi. In questo contesto, la manipolazione delle rappresentazioni informazionali (come nei deepfakes) non è fenomeno marginale ma attacco alla struttura stessa dell'esperienza vissuta.

**Intelligenza Artificiale Generale (AGI):** Ipotetico sistema di IA dotato di capacità cognitive flessibili e generali comparabili a quelle umane: apprendimento trasferibile tra domini diversi, ragionamento astratto, comprensione causale profonda, autonomia negli obiettivi. A differenza della Narrow AI, l'AGI potrebbe soddisfare i criteri tradizionali di personalità (razionalità, autonomia, intenzionalità), mettendo definitivamente in crisi il presupposto antropocentrico del diritto d'autore.

**Intenzionalità:** Proprietà degli stati mentali di essere diretti verso oggetti o stati di cose (Brentano, Husserl). Questione filosofica centrale: se l'IA possa possedere intenzionalità autentica (intrinseca, come sostengono i funzionalisti) o solo derivata/simulata (come argomenta Searle con la Stanza cinese). La risposta ha conseguenze dirette sulla possibilità di riconoscere all'IA lo status di autore.

**Intenzionalità mediata:** Prospettiva secondo cui, anche se il sistema artificiale non possiede intenzionalità propria, può veicolare l'intenzionalità dell'essere umano che lo utilizza attraverso un controllo creativo significativo sul processo generativo (specificità dei prompt, iterazione, selezione, editing).

**Latour, B. (Actor-Network Theory):** v. Actor-Network Theory.

**Licenza Creative Commons:** v. Creative Commons.

**Licensing collettivo esteso (ECL):** Meccanismo attraverso cui un ente di gestione collettiva negozia licenze con utilizzatori per conto di tutti gli autori di una categoria, inclusi non-membri, con possibilità di opt-out individuale. Operativo in Danimarca, Norvegia e Svezia per usi massivi come la radiodiffusione; proposto come modello adattabile all'addestramento di sistemi di IA.

**Machine learning:** Branchia dell'IA che sviluppa algoritmi capaci di apprendere da dati senza essere esplicitamente programmati per ogni compito specifico. Comprende supervised learning, unsupervised learning e reinforcement learning. Base tecnologica dell'IA generativa contemporanea.

**Memorizzazione (dei dati di training):** Fenomeno per cui i modelli linguistici di grandi dimensioni possono riprodurre verbatim sequenze di testo presenti nei dati di addestramento se interrogati con prompt opportunamente costruiti (Carlini et al., 2021). Implica che l'apprendimento algoritmico non è puramente

astrattivo come la lettura umana, con conseguenze per la valutazione della liceità del training su opere protette.

**Merger doctrine:** Dottrina del diritto d'autore secondo cui quando idea ed espressione sono inseparabili (merge), l'espressione non riceve protezione per evitare un monopolio sull'idea stessa. Rilevante per valutare la proteggibilità di output algoritmici strutturalmente predeterminati.

**Midjourney:** Sistema di IA generativa per immagini sviluppato dall'omonima società indipendente, accessibile tramite Discord. Tra i sistemi che hanno segnato la svolta nell'IA generativa accessibile al pubblico nel 2022, al centro di controverse giuridiche sull'addestramento su opere protette (Andersen v. Stability AI).

**Narrow AI:** Intelligenza artificiale ristretta o specializzata: sistemi eccellenti in compiti specifici (giocare a scacchi, generare immagini, tradurre testi) ma incapaci di generalizzare oltre il dominio per cui sono stati addestrati. Diversamente dall'AGI, i sistemi di IA attuali sono narrow AI.

**Neural network (Rete neurale):** Modello computazionale ispirato al funzionamento dei neuroni biologici, costituito da nodi interconnessi organizzati in livelli che trasformano input in output attraverso operazioni matematiche. Più livelli (deep neural network) consentono apprendimento di rappresentazioni sempre più astratte.

**Noema / Noesis:** Distinzione fenomenologica elaborata da Edmund Husserl: la noesis è l'atto intenzionale (il modo in cui la coscienza è diretta); il noema è l'oggetto inteso (il contenuto dell'atto intenzionale). Fondamentale nel dibattito sull'intenzionalità dell'IA: i sistemi computazionali esibiscono strutture funzionalmente analoghe alla noesis senza necessariamente avere accesso al noema.

**Oligopolio tecnologico:** Struttura di mercato in cui poche imprese (OpenAI, Google, Meta, Microsoft, Anthropic) controllano i modelli di IA più avanzati e catturano la maggior parte del valore economico generato. Derivante da barriere all'entrata enormi (costi computazionali, dataset massivi, competenze rare) e da effetti di rete. Solleva questioni di giustizia distributiva e concentrazione del potere culturale.

**Onlife:** Termine coniato da Luciano Floridi per descrivere la condizione in cui la distinzione tra vita online e offline si fa progressivamente irrilevante: gli esseri umani abitano un'unica realtà informazionale strutturalmente mediata da tecnologie digitali.

**Ontologia sociale:** Approccio filosofico che studia il diritto come sistema di fatti istituzionali ontologicamente dipendenti dall'intenzionalità collettiva (Searle). Distingue tra fatti bruti (esistenti indipendentemente dall'osservatore) e fatti istituzionali (costituiti da regole condivise). Il diritto d'autore non protegge proprietà metafisiche immanenti ma costituisce lo status giuridico dell'opera.

**Open source:** Modello di sviluppo software in cui il codice sorgente è pubblicamente accessibile, modificabile e redistribuibile. Nel contesto dell'IA generativa, il dibattito sull'open source riguarda l'accessibilità dei pesi dei modelli (open weights) e del codice, con implicazioni per la concentrazione del potere tecnologico e i rischi di usi malevoli.

**Opacità algoritmica:** Caratteristica dei sistemi di IA complessi (specialmente deep learning) i cui meccanismi interni sono difficili o impossibili da comprendere anche per i loro creatori (black box). Ostacola l'attribuzione di responsabilità, la verifica dell'originalità e l'esercizio del diritto alla spiegazione delle decisioni algoritmiche.

**Opt-out (riserva dei diritti):** Meccanismo previsto dall'art. 4(3) della Direttiva (UE) 2019/790 che consente ai titolari dei diritti di riservare espressamente i propri contenuti, escludendoli dall'eccezione generale per il text and data mining commerciale. Forma machine-readable richiesta per i contenuti online.

**Originalità:** Requisito fondamentale per la protezione del diritto d'autore: l'opera deve essere creazione intellettuale propria dell'autore, risultato di scelte libere e creative che imprimono il 'tocco personale' del creatore (standard europeo stabilito nelle sentenze Infopaq e Painer della CGUE). Nell'era dell'IA, il requisito entra in crisi perché i sistemi generativi producono novità statistica senza intenzionalità consapevole.

**Personalità giuridica:** Qualificazione normativa attraverso cui l'ordinamento decide chi riconoscere come soggetto di diritto, capace di essere titolare di situazioni giuridiche soggettive. Non è attributo naturale ma costruzione normativa storicamente variabile (estesa a schiavi e donne solo progressivamente; attribuita a società e fondazioni). Il dibattito sull'AI personhood si chiede se possa essere estesa all'IA.

**Phronesis (prudenza):** Virtù aristotelica della saggezza pratica: capacità di deliberare bene circa i mezzi appropriati per raggiungere fini buoni in situazioni concrete di incertezza. Nel contesto dell'IA generativa, virtù essenziale per progettisti e decisori che devono valutare rischi, prevedere usi impropri e bilanciare innovazione e precauzione.

**Poiesis / Praxis:** Distinzione aristotelica tra produzione orientata all'opera esterna (poiesis) e azione che trova in sé il proprio fine (praxis). Rilevante per distinguere la produzione algoritmica (assimilabile alla poiesis meccanica) dalla creazione umana autentica (che incorpora elementi di praxis).

**Principio di differenza (Rawls):** Secondo principio di giustizia rawlsiano: le disuguaglianze socioeconomiche sono giustificate solo se massimizzano le aspettative dei membri meno avvantaggiati della società. Applicato all'economia dell'IA, richiede che la protezione delle innovazioni sia giustificata solo se agevola i più svantaggiati, non se serve esclusivamente ad arricchire élite tecnologiche.

**Prompt:** Input testuale (o multimodale) fornito dall'utilizzatore a un sistema di IA generativa per guidare la produzione dell'output desiderato. La specificità e complessità del prompt è uno dei criteri per valutare il controllo creativo significativo esercitato dall'utilizzatore ai fini dell'attribuzione dell'autorialità.

**Prompt engineering:** Pratica di progettazione e ottimizzazione dei prompt per ottenere output di qualità superiore da sistemi di IA generativa. Richiede competenze tecniche e creative specifiche; può costituire contributo creativo rilevante ai fini dell'attribuzione dei diritti d'autore.

**Qualia:** Aspetti qualitativi soggettivi dell'esperienza cosciente: 'che cosa si prova' a vedere il rosso, ascoltare musica, provare dolore. Il problema difficile della coscienza (Chalmers) riguarda perché le funzioni cognitive siano accompagnate da qualia. Rilevante per il diritto d'autore: se la tutela presuppone il vissuto interiore del processo creativo, l'assenza di qualia nell'IA può giustificare un trattamento diversificato.

**Qubit:** Unità fondamentale di informazione nell'informatica quantistica. A differenza del bit classico (0 o 1), il qubit può esistere in sovrapposizione di stati, sfruttare l'entanglement e l'interferenza quantistica per risolvere classi di problemi intrattabili su hardware classico.

**Recombinantics:** Termine tecnico coniato da David Cope per descrivere il processo attraverso cui il programma EMI (Experiments in Musical Intelligence) riutilizza e ricombina segmenti di opere musicali esistenti. Esempifica una concezione funzionale e comportamentale della creatività algoritmica come ricombinazione di pattern.

**Reflective equilibrium (Equilibrio riflessivo):** Metodo di riflessione morale elaborato da John Rawls: si parte dai principi morali più solidi e si aggiustano reciprocamente alla luce delle conseguenze concrete, cercando coerenza tra intuizioni particolari e principi generali. Adottato come approccio pluralista nella governance dell'IA: i doveri deontologici fissano i vincoli, il calcolo delle conseguenze guida le scelte tra opzioni che rispettano quei vincoli.

**Regulatory sandbox:** Ambiente controllato in cui imprese possono testare prodotti o servizi innovativi beneficiando di esenzioni temporanee da alcune norme regolatorie, sotto supervisione dell'autorità competente. Previsto dagli artt. 57-63 dell'AI Act (obbligo per gli Stati membri di istituirli entro agosto 2026). Strumento per bilanciare innovazione e protezione dei diritti.

**Reinforcement learning:** Paradigma di machine learning in cui un agente apprende attraverso interazione con un ambiente, ricevendo ricompense o penalità in base alle azioni compiute. I sistemi di reinforcement learning possono evolvere autonomamente senza supervisione umana diretta, complicando l'attribuzione della paternità degli output.

**Singularità tecnologica:** Ipotetico punto di svolta tecnologico in cui l'IA supererebbe l'intelligenza umana e si autopotenzierebbe in modo accelerato, rendendo impossibile prevedere gli sviluppi successivi. Secondo Kurzweil, si collocherebbe intorno al 2045. Il concetto mette in discussione i presupposti antropocentrici del diritto e della filosofia.

**Soft law:** Strumenti normativi privi di forza cogente ma dotati di autorevolezza morale e funzione orientativa: codici di condotta, linee guida, raccomandazioni. Nel contesto dell'AI Act, la distinzione tra codici di pratica per modelli GPAI (art. 56) e codici di condotta volontari (art. 95) articola diversi livelli di soft law integrata nel framework normativo vincolante.

**Soggettività giuridica:** Capacità di essere titolari di situazioni giuridiche soggettive (diritti e obblighi). Tradizionalmente riservata a persone fisiche e giuridiche. Il dibattito sull'AI personhood o electronic personhood esplora l'estensione di questa capacità a sistemi artificiali, con implicazioni profonde per il diritto d'autore, la responsabilità e i diritti fondamentali.

**Spazio latente:** Rappresentazione matematica multidimensionale appresa da un modello generativo durante l'addestramento, che codifica le strutture statistiche e semantiche dei dati. La generazione di nuovi contenuti avviene per interpolazione nello spazio latente, producendo output che non esistevano nel corpus originale.

**Stanza cinese (esperimento mentale):** Esperimento mentale elaborato da John Searle: una persona che non comprende il cinese chiusa in una stanza con un manuale di regole sintattiche produce risposte apparentemente intelligenti senza comprensione semantica. Argomenta che la manipolazione sintattica di simboli (come nell'IA) non equivale a comprensione semantica autentica, dunque l'IA non possiede intenzionalità intrinseca.

**Statuto di Anna (1710):** Prima legge moderna sul copyright, approvata in Gran Bretagna nel 1710. Attribuiva per la prima volta agli autori (e non agli editori) il diritto esclusivo di stampare e vendere le proprie opere per un periodo determinato. Fondamento storico della concezione moderna dell'autorialità.

**Sui generis:** Espressione latina ('di genere proprio') che designa forme di tutela giuridica specifiche create per fenomeni che non rientrano nelle categorie tradizionali. Nel contesto dell'IA, la tutela sui generis per le opere generate algoritmicamente (ispirata ai diritti sulle banche dati della Direttiva 96/9/CE) rappresenta un'alternativa alla piena autorialità e al dominio pubblico.

**Tertium genus:** Categoria concettuale che designa l'IA generativa come fenomeno genuinamente nuovo: né mero strumento passivo né autore pienamente autonomo, ma entità che esibisce forme di agency limitata, produce output originali attraverso processi deterministici e stocastici e opera in spazi concettuali che possono trascendere la comprensione umana diretta.

**Text and data mining (TDM):** Analisi automatizzata di grandi quantità di testi o dati per estrarre pattern, informazioni o conoscenze. La Direttiva (UE) 2019/790 prevede un'eccezione obbligatoria per il TDM a fini di ricerca scientifica (art. 3) e un'eccezione generale per qualsiasi finalità con possibilità di opt-out per i titolari (art. 4). L'applicabilità di questo quadro all'addestramento di modelli commerciali rimane controversa.

**Transformer:** Architettura di rete neurale basata su meccanismi di attenzione che ha rivoluzionato il natural language processing e l'IA generativa. Introdotta nel 2017 (Vaswani et al., 'Attention Is All You Need'), costituisce la base di modelli come GPT, BERT, LLaMA e dei principali sistemi generativi attuali.

**Transformative use:** Elemento chiave della dottrina del fair use statunitense: un uso è trasformativo quando aggiunge nuovi significati, espressioni o messaggi all'opera originale, servendo una funzione diversa. La Corte Suprema in *Andy Warhol Foundation v. Goldsmith* (2023) ha ridimensionato la portata del concetto, chiarendo che la presenza di un nuovo 'messaggio' non basta da sola a giustificare il fair use.

**Tutela sui generis:** v. Sui generis.

**Universal Basic Income (UBI):** Proposta di reddito di base universale finanziato attraverso tassazione delle tecnologie che automatizzano il lavoro umano, incluso quello creativo. Nel contesto dell'IA generativa, consentirebbe a creativi di non dipendere dal mercato per la propria sussistenza. Dibattuta per fattibilità fiscale e effetti macroeconomici.

**Urheberrecht:** Sistema tedesco del diritto d'autore, espressione paradigmatica della tradizione continentale personalistica: l'opera è indissolubilmente legata alla personalità dell'autore e i diritti morali sono inalienabili. Fortemente influenzato dalla teoria hegeliana della proprietà come oggettivazione della personalità.

**Utilitarismo:** Teoria etica secondo cui la giustezza di un'azione dipende dalle sue conseguenze in termini di massimizzazione del benessere complessivo (Bentham, Mill). Applicato al diritto d'autore, giustifica la protezione come strumento per incentivare la produzione di opere socialmente utili (Copyright Clause della Costituzione USA). Nell'era dell'IA, pone la questione se la protezione delle opere algoritmiche massimizzi davvero il benessere collettivo.

**Velo di ignoranza (Rawls):** Dispositivo procedurale elaborato da John Rawls: i principi di giustizia vengono scelti da individui razionali che non conoscono la propria posizione nella società (classe, talento, genere). Assicura imparzialità nella selezione dei principi. Applicato all'economia dell'IA, suggerisce che i principi regolativi dovrebbero essere scelti senza sapere se si sarà autori, sviluppatori o utenti finali.

**Watermarking:** Tecnica di marcatura digitale invisibile inserita negli output dei sistemi di IA generativa per identificarne l'origine artificiale. Strumento di trasparenza e tracciabilità previsto tra gli obblighi di accountability proposti per i provider di sistemi generativi.

**Zombie filosofico (Chalmers):** Essere ipoteticamente identico a un essere umano sul piano funzionale e comportamentale ma privo di esperienza soggettiva (qualia). Argomento elaborato da David Chalmers per sostenere l'irriducibilità della coscienza fenomenica a spiegazioni puramente fisiche o funzionali. Applicato all'IA: anche un sistema perfettamente equivalente all'umano sul piano computazionale potrebbe difettare di coscienza, con conseguenze per il riconoscimento dell'autorialità.



Finito di stampare nel mese di maggio 2026  
dalla *Grafica Elettronica* - Napoli