

# Opacità algoritmica e sovranità epistemica nel contesto del capitalismo delle piattaforme

Piero Sansò<sup>1 2 3</sup>

<sup>1</sup> Università di Modena e Reggio Emilia

<sup>2</sup> Fondazione Collegio San Carlo

<sup>3</sup> Almo Collegio Borromeo

**Abstract:** Il contributo esamina i complessi rapporti tra opacità algoritmica e controllo epistemico nelle società digitali contemporanee, con particolare riferimento alle dinamiche del *Surveillance Capitalism*. L'analisi si concentra sulla tensione tra l'efficienza computazionale dei sistemi di intelligenza artificiale basati su *machine learning* e il perimetro di opacità che spesso circonda queste tecnologie. Riprendendo una concezione dell'opacità che la vede riferita anche ai fattori non eminentemente tecnici – quali segretezza del codice e analfabetismo digitale – l'articolo analizza le modalità con cui essa si configura quale dispositivo di potere epistemico. Infine, si sottolinea come, oltre ai ben noti rischi per i diritti soggettivi, la *black box* algoritmica presenti delle criticità sistemiche rispetto alle categorie di sovranità e agentività epistemica, intese come necessarie allo sviluppo di processi democratici sani e autentici.

**Keywords:** Intelligenza artificiale, opacità, trasparenza, democrazia, sovranità epistemica

## 1 Introduzione: il problema della *black box*

La diffusione capillare di sistemi di intelligenza artificiale basati sull'apprendimento autonomo a partire dai dati non si limita a fornire nuovi strumenti tecnici: essa pone interrogativi e sfide che attraversano in maniera ubiqua i campi del sapere e dell'azione umana. La storia dell'intelligenza artificiale, fatta tanto di “estati” quanto di “inverni”, sembra essere giunta a una svolta di natura epocale<sup>1</sup>. Negli ultimi due decenni siffatte tecnologie hanno infatti assunto una funzione infrastrutturale nei processi decisionali pubblici e privati, dalla concessione di credito alla selezione del personale, dal governo dei flussi informativi online alla distribuzione di servizi essenziali. La rapidità e la pervasività di tale diffusione ne hanno imposto la centralità anche nel dibattito filosofico e giuridico, sia internazionale sia italiano<sup>2</sup>. Per comprendere appieno le implicazioni di tale

✉ [piero.sanso@unimore.it](mailto:piero.sanso@unimore.it) (Piero Sansò);

1. Floridi, Luciano. *La quarta rivoluzione industriale. Come l'infosfera sta trasformando il mondo*, Cortina, Milano, 2017
2. Vedasi, fra gli altri, Amato, Sabrina. *Biodiritto 4.0. Intelligenza artificiale e nuove tecnologie*. Giappichelli, Torino, 2020; Casadei, Thomas e Pietropaoli, Stefano. “Intelligenza artificiale: l'ultima sfida per il diritto?” *Diritto e tecnologie informatiche. Questioni di informatica giuridica, prospettive istituzionali e sfide sociali*. A cura di Th. Casadei and S. Pietropaoli. Wolters Kluwer, Milano, 2024. 259–274; Durante, Massimo. *Potere computazionale. L'impatto delle ICT su diritto, società, sapere*. Mimesis, Milano, 2019; Floridi, Luciano. *Etica dell'intelligenza artificiale, sviluppi, opportunità, sfide*. Raffaello Cortina Editore, Milano, 2022; Oliveri, Federico. *Machina Mundi. Per una regolazione democratica dei poteri digitali*, Mucchi editore, Modena, 2025; Palazzani, Laura. *Tecnologie dell'informazione e intelligenza artificiale. Sfide etiche al diritto*. Studium, Roma, 2020; Sapienza, Salvatore. *Decisioni algoritmiche e diritto*. Giuffrè Francis Lefebvre, Milano, 2024; Sartor, Giovanni. *L'intelligenza artificiale e il diritto*. Giappichelli, Torino, 2022.

pervasività tecnologica, è necessario precisare i diversi approcci metodologici che caratterizzano i sistemi di IA, ciascuno dei quali presenta specifiche caratteristiche in termini di trasparenza e interpretabilità.

Allo stato attuale, i modelli di rappresentazione formale della conoscenza si articolano essenzialmente attorno a due paradigmi fondamentali. Il primo è il paradigma procedurale, consolidatosi nell'alveo dell'informatica novecentesca, secondo il quale le conoscenze vengono formulate in sequenze operative che vengono impartite al calcolatore: in tal modo, il sapere risulta incorporato nelle stesse operazioni computazionali che l'algoritmo è chiamato ad eseguire<sup>3</sup>. I sistemi che adottano un approccio procedurale sono caratterizzati dalla capacità di restituire esiti univoci e prevedibili, in quanto la loro attività consiste nell'eseguire fedelmente sequenze di istruzioni (subroutine o funzioni) predefinite. I sistemi di intelligenza artificiale simbolica – la cosiddetta *Good Old Fashioned Artificial Intelligence*<sup>4</sup> – quali, ad esempio i sistemi esperti, ricadono in questa categoria. Tuttavia, l'evoluzione tecnologica degli ultimi decenni ha evidenziato i limiti di questo approccio, particolarmente quando applicato a domini complessi caratterizzati da incertezza e ambiguità. A partire dallo sviluppo della rete e dalla circolazione di enormi quantità di dati digitali, è riemerso un progetto alternativo che ribalta radicalmente la logica tradizionale della programmazione, quello connessionista-dichiarativo.

In contrapposizione rispetto alle logiche procedurali, il paradigma dichiarativo prospetta una separazione tra le informazioni inerenti al problema da risolvere e il procedimento di elaborazione: in questa prospettiva la risoluzione di problemi specifici viene affidata a un motore inferenziale che genera nuova conoscenza. Le odierne tecnologie di intelligenza artificiale si ancorano prevalentemente a questo secondo modello, offrendo una rappresentazione subsimbolica della conoscenza e dei problemi che non si configura dunque come rigidamente deterministica<sup>5</sup>. Rientrano in questo secondo insieme i più diffusi, ad oggi, sistemi di intelligenza artificiale (*machine e deep learning, artificial neural network*).

Questa distinzione metodologica tra paradigmi simbolici e subsimbolici assume un'importanza cruciale quando si analizza una delle problematiche centrali emerse dalla più recente stagione dell'intelligenza artificiale: la tensione tra efficienza computazionale e trasparenza decisionale, che si manifesta nell'opposizione dicotomica fra opacità e trasparenza. Difatti, i modelli simbolici, basati sulla rappresentazione formale, permettono un'individuazione precisa dei processi interni che generano l'output. Al contrario, i modelli subsimbolici, basati sull'apprendimento a partire dai dati, non permettono di tracciare con facilità e precisione i propri percorsi di ragionamento, a causa delle caratteristiche interne all'architettura – non determinata *ex ante*, ma appresa – che li compone. Spesso, dunque, la maniera in cui i sistemi giungono agli esiti prodotti restano non completamente decifrabili neppure dagli sviluppatori, o analizzando codice sorgente, dataset e documentazione tecnica. La differente tecnologia alla base di questi modelli incide direttamente sulla capacità di spiegarne il funzionamento, ponendo dunque la problematica della decisione, sia individuale sia istituzionale, a partire dai processi opachi e automatici tramite cui i risultati vengono generati.

La natura opaca di siffatti sistemi di IA è stata metaforicamente resa, da una decina di anni a questa parte, nei termini di una *black box*<sup>6</sup>, una scatola nera. Si definisce scatola nera un processo di cui sono unicamente visibili input e output, mentre restano opachi i meccanismi interni. A ciò si legano le esigenze di trasparenza invocate sia eticamente che normativamente. Tuttavia, è bene specificare, nell'ambito dell'IA, di quale specifica idea di opacità si voglia impiegare.

L'opacità può essere considerata, seguendo la definizione di Humphreys, nei seguenti termini<sup>7</sup>. Un processo è epistemicamente opaco relativamente a un agente cognitivo X al tempo t solo nel caso in cui X non conosca al tempo t tutti gli elementi epistemicamente rilevanti del processo. Questa formulazione è stata chiamata anche

3. Cfr. Pietropaoli, Stefano. "Verso un legislatore non umano? Brevi riflessioni su alcuni problemi di diritto computazionale" *Osservatorio sulle fonti*, 2(2022): 397-410

4. Per una panoramica sul tema, si veda Boden, Margaret A. "GOFAI." *The Cambridge Handbook of Artificial Intelligence*. Keith Frankish and William M. Ramsey. Cambridge (a cura di). Cambridge University Press, Cambridge, 2014: 89-107.

5. Cfr. Pietropaoli, Stefano, *op. cit.*, pp. 402-403.

6. A rendere popolare il termine in relazione all'intelligenza artificiale ha enormemente contribuito Pasquale, Frank. *The Black Box Society: The Secret Algorithms That Control Money and Information*. Harvard University Press, Cambridge/Londra, 2015.

7. Humphreys, Paul. "The Philosophical Novelty of Computer Simulation Methods." *Synthese* 169.3 (2009): 615-626. In particolare, cfr. p. 618.

“opacità epistemica debole”<sup>8</sup>, essendo contrapposta a un’opacità essenziale, definibile nei seguenti termini. Un processo è essenzialmente epistemicamente opaco per un agente cognitivo X se e solo se è impossibile, data la natura di X, che X conosca tutti gli elementi epistemicamente rilevanti del processo<sup>9</sup>. Applicare questa seconda definizione, significa spostare interamente l’attenzione sulle caratteristiche strutturali dei processi, lasciando fuori ogni tipo di elemento di natura contestuale e concentrandosi sulla composizione strutturale dell’oggetto osservato.

Al contrario, l’opacità intesa nella sua forma “debole”, designa un campo epistemico retto da tre elementi: l’agente cognitivo X, il processo interessato e il tempo t, lo specifico momento temporale in cui avviene l’interazione fra soggetto e oggetto. Assumono dunque fondamentale rilevanza non solo le caratteristiche del processo osservato, ma la dimensione storica, soggettiva e contestuale, in cui l’atto conoscitivo si inserisce. Così intesa l’opacità dipende strutturalmente dalle forme di organizzazione economica e politica dello sviluppo e distribuzione della tecnologia, abilitando una riflessione critica che coinvolga anche la dimensione storica e sociale degli artefatti. Questa articolazione osserva dunque l’opacità algoritmica non quale fenomeno monolitico, ma piuttosto una stratificata rete di istanze opache fra loro interdipendenti. La natura relazionale e contestuale di siffatta interpretazione dei processi opachi apre lo spazio per programmi di intervento mirato volti a rendere trasparenti singoli nodi della “rete” opacizzante; anche nell’impossibilità di una trasparenza e spiegabilità radicali legate al modello. Solo attraverso questa consapevolezza teorica sono possibili operazioni di critica filosofica, economica e giuridica del fenomeno dell’IA, che portino infine a disegnare programmi culturali e normativi efficaci. Questo contributo muove da tale presupposto, considerando i sistemi di intelligenza artificiale e l’opacità che in questa fase storica li caratterizza quali sistemi socio-tecnici, che incorporano dunque in sé presupposti e ragioni socialmente e soggettivamente determinate<sup>10</sup>. Si argomenterà dunque in supporto della tesi che vede l’opacità algoritmica, nelle forme storicamente situate in cui la si sta sperimentando, costituire un rischio sistemico inedito per l’agentività epistemica dei cittadini e la sovranità epistemica dei processi democratici.

## 2 Opacità algoritmica: dimensioni strutturale e socialmente costituita

Per affrontare in modo mirato il problema della *black box* algoritmica, è dunque utile distinguere due macro-categorie di opacità: da un lato un’opacità di natura *strutturale, interna* al funzionamento tecnico dei sistemi di IA; dall’altro un’opacità *socialmente costituita, esterna* ai sistemi, derivante dai fattori umani, sociali e valoriali che determinano gli spazi di produzione, implementazione e utilizzo degli stessi<sup>11</sup>. Come visto, l’opacità può dunque essere intesa anche nei termini di fenomeno situato e multi-agente<sup>12</sup>, non unicamente dipendente dalla struttura di apprendimento interna del sistema, ma anche da condizioni materiali, dalla disponibilità o meno del codice e dei file di log<sup>13</sup>, dall’ambito di applicazione, dalle competenze e dagli scopi degli attori

8. Per un approfondimento sul tam: San Pedro García, Inaki. “Degrees of Epistemic Opacity”. *Teorema. International Journal of Philosophy*, 43.2(2024): 5-21

9. Humphreys, Paul, *op. cit.* p. 618.

10. Su queste tematiche è indispensabile citare, Zuboff, Shoshana. *Il capitalismo della sorveglianza. Il futuro dell’umanità nell’era dei nuovi poteri*. LUISS University Press, 2020

11. Burrell, J. “How the machine ‘thinks’: Understanding opacity in machine learning algorithms”. *Big Data & Society*, 3.1(2016). L’autrice, in una delle prime formalizzazioni tassonomiche del problema, individua tre forme di opacità: “opacity as intentional corporate or state secrecy” legata al segreto industriale, alla proprietà intellettuale, e alla protezione dei vantaggi competitivi; “opacity as technical illiteracy” e opacity that arises from the characteristics of machine learning algorithms”, che si costituisce a livello logico e semantico nei sistemi, anche a causa delle dimensioni che essi richiedono per massimizzare l’efficienza.

12. Interessante in questo senso la proposta presente in Christin, Angèle. “The ethnographer and the algorithm: beyond the black box.” *Theory and Society* 49 (2020): 897-918. L’autrice propone un approccio sociologico e etnografico al problema della *black box*, articolando tre strategie ermeneutiche. La prima, *algorithmic refraction*, si concentra sulle riconfigurazioni che avvengono quando software computazionali, persone e istituzioni interagiscono tra loro. La seconda strategia, *algorithmic comparison*, si basa su un approccio di similarità e differenza per individuare le caratteristiche uniche degli strumenti analizzati. La terza strategia, *algorithmic triangulation*, utilizza gli algoritmi per raccogliere dati qualitativi approfonditi.

13. Si sottolinea come questa componente sia probabilmente quella fondamentale. Pensare utopistici interventi giuridici che impongano trasparenza totale e radicale del codice porta tuttavia con sé altrettante criticità. Su questi profili vedasi Fioriglio, Gianluigi. “La Società algoritmica fra opacità e spiegabilità: profili informatico-giuridici”, in *Ars interpretandi, Rivista di ermeneutica giuridica*.

umani coinvolti nella progettazione, commercializzazione e utilizzo della tecnologia.

Questo indirizzo interpretativo, emerso in differenti articolazioni all'interno della letteratura in materia<sup>14</sup>, ha anche l'obiettivo fattuale di abilitare l'elaborazione strumenti giuridici diversificati e specifici a seconda della causa dell'opacità; piuttosto che fare affidamento su di una generica nozione di trasparenza che – allo stato attuale – appare non in grado di trasformarsi da principio in prassi, limitando di conseguenza la portata concreta dei tentativi di una regolamentazione sostanziale in materia di intelligenza artificiale e, più in generale, dei più frequentati ambienti digitali (quali *social network*, *browser* e siti di *e-commerce*). Ciò implica che al fine di raggiungere l'obiettivo di un'IA autenticamente più “trasparente”, sia necessario elaborare strumenti culturali e normativi che rispondano alla sfera socialmente costituita dell'opacità, senza potersi limitare unicamente a soluzioni – comunque necessarie – di carattere tecnico, quali lo sviluppo di modelli trasparenti *by design*<sup>15</sup> e sistemi di spiegazione *post hoc*.<sup>16</sup>

L'opacità interna attiene dunque alla complessità tecnica intrinseca dei sistemi di intelligenza artificiale basati sull'apprendimento automatico a partire dai dati, parliamo dunque di tecnologie quali il *machine learning*, il *deep learning* e le *deep neural network*<sup>17</sup>. Come si è visto, questa forma di opacità non caratterizza al contrario i modelli, ad esempio i sistemi esperti, basati invece su rappresentazioni simboliche e logica procedurale. Al contrario, i modelli basati sull'apprendimento operano in architetture multilivello estremamente complesse, giungendo ai risultati tramite correlazioni statistiche, piuttosto che tramite elaborazioni deduttive. Nelle reti neurali artificiali, ad esempio, gli output sono il risultato dell'attivazione di un enorme numero di neuroni artificiali disposti su molteplici livelli interconnessi: la scala quantitativa dei parametri coinvolti costituisce il primo, estremamente significativo, motivo di difficoltà interpretativa. In aggiunta, riprendendo le riflessioni di Boge, è possibile individuare due distinte dimensioni - potenzialmente interconnesse nella pratica, ma logicamente indipendenti - del problema della *black box* nei sistemi di deep learning, quelle che l'autore chiama *how opacity* e *what opacity*<sup>18</sup>. La prima riguarda l'incapacità di comprendere in dettaglio come un modello di rete modifichi e aggiorni internamente i suoi parametri e le sue connessioni durante la fase di apprendimento, il processo algoritmico e computazionale che porta il modello a ottenere previsioni accurate. La seconda, invece, si riferisce all'incertezza e all'ignoranza riguardo a quali caratteristiche o pattern specifici la rete abbia automaticamente individuato e sfruttato per raggiungere la propria capacità predittiva. Queste due forme di opacità sono indipendenti in quanto possono essere affrontate separatamente dalle tecniche di spiegazione

---

1(2021), pp. 53-67.

14. In questo senso, appare particolarmente valida e dettagliata la tassonomia proposta in Facchini, Alessandro, and Termine, Alberto. "Towards a Taxonomy for the Opacity of AI Systems." In: Müller, Vincent C. (ed.) *Philosophy and Theory of Artificial Intelligence 2021. Studies in Applied Philosophy, Epistemology and Rational Ethics 63* (2022): Springer., Cham. Facchini e Termine elaborano una tassonomia articolata su tre distinti livelli logicamente indipendenti (l'occorrenza di uno di essi non rappresenta condizione né necessaria né sufficiente per l'occorrenza di un altro). L' *Access Opacity*: si riferisce ai limiti epistemici nel comprendere struttura e funzionamento di un sistema AI. Può derivare da tre fattori: policy di non-disclosure adottate dai progettisti; insufficiente alfabetizzazione tecnica dei soggetti coinvolti; complessità intrinseca dell'architettura computazionale. La *Link Opacity*: si riferisce all'incapacità del modello appreso di fornire elementi esplicativi, predittivi e di controllo riguardo al fenomeno oggetto di modellizzazione. Tale opacità compromette la comprensione scientifica e si suddivide in tre forme: *causal opacity*, relativa all'incapacità di ricostruire catene causali; *mechanisms opacity*, ovvero l'impossibilità di ipotizzare meccanismi sottostanti al fenomeno; *laws opacity*, l'incapacità di inferire leggi strutturali che connettano osservabili e variabili latenti. La *Semantic Opacity*: concerne l'opacità semantica dei contenuti informativi appresi dal sistema, ovvero la difficoltà, per i soggetti umani, di attribuire un significato comprensibile alle informazioni immagazzinate o manipolate. Essa si divide in: (a) *content opacity*, dovuta al formato strutturale dei dati (es. vettori numerici senza mappatura semantica); (b) *inferential opacity*, derivante dalla natura dei processi inferenziali interni al sistema (ad es. calcoli numerici non riconducibili a inferenze logiche esplicite).
15. Per un confronto fra queste due strategie, orientato verso i vantaggi della trasparenza by design, vedasi Felzmann, Helke et al. "Towards Transparency by Design for Artificial Intelligence." *Science and Engineering Ethics*, 26 (2020): 3333 – 3361 e Loi, Michele et al. "Transparency as design publicity: explaining and justifying inscrutable algorithms." *Ethics and Information Technology*, 23 (2019): 253 – 263.
16. Per una prima panoramica del campo dell' *Explainable AI* e delle principali tecniche e problematiche della spiegazione *post hoc*, vedasi Sajid Ali et al. "Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence." *Inf. Fusion*, 99 (2023): 101805.
17. Per una panoramica sul tema, si veda Boden, Margaret A. "GOFAL." *The Cambridge Handbook of Artificial Intelligence*. Ed. Keith Frankish and William M. Ramsey. Cambridge: Cambridge University Press, 2014. 89–107. Print.
18. Boge, Florian J. "Two Dimensions of Opacity and the Deep Learning Predicament." *Minds and Machines* 32.1 (2022): 43-75.

*post hoc*, utilizzando metodi e strumenti che fanno riferimento a variabili e aspetti completamente diversi del processo decisionale del sistema algoritmico<sup>19</sup>.

Se l'opacità interna trova la propria ragione nelle complessità strutturali e nei limiti epistemici dei sistemi di apprendimento automatico, il quadro diviene ancor più articolato quando si considerano le forme di opacità che emergono dalla matrice sociale e istituzionale entro cui tali sistemi vengono pensati, sviluppati e commercializzati. In questa prospettiva, l'opacità non appare come un fenomeno esclusivamente riconducibile al dominio del tecnico, ma come esito di una complessa stratificazione di saperi, pratiche e poteri che travalicano la dimensione puramente algoritmica.

Una dimensione centrale di questa opacità socialmente costituita attiene al deficit di competenze, ma soprattutto di consapevolezza, della cittadinanza rispetto al riconoscimento e all'interazione attiva con l'IA. Quello che Burrell definisce "technical illiteracy"<sup>20</sup> si estende significativamente anche a categorie professionali che operano – o che si pianifica di far operare a breve - a stretto contatto con sistemi di IA, ponendo la questione dell'educazione al digitale in una posizione di assoluta centralità<sup>21</sup>. L'alfabetizzazione algoritmica – intesa come capacità di riconoscere la presenza, comprendere le logiche operative e valutare criticamente l'impatto dei sistemi algoritmici nella propria esperienza lavorativa e non – sta divenendo prerequisito essenziale per l'esercizio consapevole della cittadinanza nell'era digitale<sup>22</sup>. L'attuale asimmetria fra diffusione di processi algoritmici e consapevolezza rispetto agli stessi crea un contesto in cui la maggioranza degli utenti si trova esposta agli effetti dei sistemi "intelligenti" senza possedere gli strumenti conoscitivi per comprenderli, contestarli o indirizzarli verso le finalità desiderate<sup>23</sup>.

Altra fonte di opacità è la condizione di dipendenza epistemica che caratterizza le comunità scientifiche e professionali coinvolte nello sviluppo e nella gestione dei sistemi intelligenti. La crescente specializzazione rende irrealistico che un singolo individuo, per quanto specializzato, possa comprendere con pari competenza di tutte le componenti rilevanti di un sistema complesso. La conoscenza di essi diviene dunque necessariamente distribuita: ciascun attore partecipa, con il proprio sapere specialistico, al funzionamento complessivo del sistema, ma nessuno ne detiene un dominio assoluto. In questa realtà, la conoscenza del singolo è inevitabilmente mediata dalla fiducia riposta nell'*expertise* altrui – una fiducia che, lungi dall'essere mero accidente, diviene condizione strutturale della pratica scientifica e tecnica contemporanea.<sup>24</sup>

Infine, accanto all'opacità non intenzionale si colloca quella frutto di strategie deliberate di riservatezza. Codice sorgente, documentazione tecnica e tracciamento dei log sono molto spesso inaccessibili, salvo interventi delle autorità giudiziali. La protezione del *know-how* aziendale, il ricorso a clausole di segretezza e la tutela della proprietà intellettuale costituiscono una delle principali cause – se non la principale - del perimetro di opacità intorno ai sistemi algoritmici. Ciò è sia vero per quanto riguarda sistemi "specialistici" – si pensi ai modelli di IA impiegati a supporto della diagnosi medica<sup>25</sup> o a quelli talvolta usati dalla pubblica

19. Ivi, pp. 64-69.

20. Burrell, Jane, *op. cit.*

21. Sul tema dell'educazione al diritto nel contesto delle nuove tecnologie, Casadei, Thomas; Marzocco Valeria e Zullo Silvia. *La didattica del diritto. Metodi, strumenti e prospettive* (seconda edizione). Pancini Editore, Pisa, 2021.

22. Per una ricognizione, non esaustiva, su un tema estremamente vasto, si vedano: Brighi, Raffaella e Zullo, Silvia. *Filosofia del diritto e nuove tecnologie. Prospettive di ricerca tra teoria e pratica*. Aracne, Roma, 2015; Pascuzzi, Giovanni. *La cittadinanza digitale. Competenze, diritti e regole per vivere in rete*. Il Mulino, Bologna, 2021; Scagliarini, Simone. "I diritti costituzionali nell'era di internet: cittadinanza digitale, accesso alla rete e net neutrality". In Casadei, Thomas e Pietropaoli, Stefano (a cura di) *op. cit.* pp 3-13

23. Sul tema si veda Martoni, Michele. "Un'autonomia ostacolata: Limiti cognitivi, incompetenze e design ingannevoli nella trasformazione digitale." *Sociologia del diritto*, 51.1(2024): 7-32. Nello specifico, pp. 15-18.

24. Mann, Sara, et al. "Sources of Opacity in Computer Systems: Towards a Comprehensive Taxonomy." *IEEE 31st International Requirements Engineering Conference Workshops (REW)*, 2023: 334-344.

25. Sul rapporto che intercorre fra pratica medica, informatica di settore e diritto si suggerisce Fioriglio, Gianluigi. *Informatica medica e diritto. Un'introduzione*, Mucchi, Modena 2020. Sulle criticità che l'opacità algoritmica pone in ambito medico e diagnostico e sugli strumenti per superarle si vedano anche Winter, Peter D. e Carusi Annamaria "(De)troubling transparency: artificial intelligence (AI) for clinical applications" *Medical Humanities* 49(2023):17-26 e Schmidt, Eva et al. "The Epistemic Cost of Opacity: How the Use of Artificial Intelligence Undermines the Knowledge of Medical Doctors in High-Stakes Contexts". *Philos. Technol.* 38.5 (2025).

amministrazione<sup>26</sup> – sia per quanto riguarda le architetture algoritmiche che strutturano le logiche delle principali piattaforme digitali, ad esempio quelle di raccomandazione dei contenuti e delle pubblicità.

L'opacità, dunque, non si lascia ricondurre a una dimensione solo tecnica o solo sociale, ma nasce nel punto di intersezione tra le condizioni materiali della progettazione, le necessità istituzionali, le strategie economiche e il tessuto relazionale delle competenze e delle responsabilità. La filosofia del diritto e l'informatica giuridica hanno la potenzialità, in questo senso, di offrire quello sguardo d'insieme necessario all'inquadramento e alla mitigazione delle criticità in materia. Il problema della trasparenza algoritmica non può essere ridotto né a una questione di accessibilità tecnica né a una battaglia formale per la *disclosure* normativa, ma richiede una riflessione sullo statuto stesso del sapere, sulla divisione sociale della competenza e sulle forme istituzionali della fiducia e del controllo

Tale distinzione fra le diverse forme di opacità non è meramente teorica, ma ha ricadute regolative rilevanti: una regolazione informata dell'IA dovrà dunque riconoscerne la natura sfaccettata, contrastandone sia le cause tecniche che quelle sociali. Sebbene, in linea teorica, sia questa la direzione presa dal Regolamento (UE) 2024/1689 del Parlamento europeo (13 giugno 2024), il documento presenta tuttavia esso stesso un certo livello di opacità, articolate attorno a criticità interpretative fondamentali che potrebbe infine contribuire a minarne l'efficacia<sup>27</sup>. Rilevante per i temi della presente trattazione è la categoria della trasparenza che, per quanto centrale, non appare definita in maniera precisa. In una recente proposta ermeneutica, Buttaboni e Floridi<sup>28</sup> propongono al contrario di suddividerla in quattro ordini, concettualmente distinti ma dipendenti l'uno dall'altro. Il primo è la trasparenza, vale a dire la disponibilità *ex ante* di documentazione tecnica, dati di training e finalità dichiarate del sistema; il secondo è la tracciabilità, ossia la conservazione di log, versioning e data lineage capaci di ricostruire *ex post* il processo decisionale, imponendo al titolare di dimostrare la conformità del trattamento. Il terzo livello è l'interpretabilità: riguarda la possibilità, per un soggetto tecnicamente competente, di comprendere la logica di funzionamento del modello (si pensi a surrogate models o visualizzazioni delle feature). Infine, il quarto livello è la spiegabilità, intesa come giustificazione teleologica rivolta a un destinatario determinato: condizione necessaria per consentire contestazione e controllo democratico.

L'adozione di una simile tassonomia permette di pensare l'opacità quale fenomeno organizzato su diversi piani reciprocamente interdipendenti. La mancanza di trasparenza a un livello inferiore propaga opacità che non vengono risolte dall'adempimento alle norme di un altro livello: senza trasparenza non c'è tracciabilità, senza tracciabilità l'interpretabilità vede cadere i suoi presupposti e, in mancanza di interpretabilità, la spiegabilità rischia di trasformarsi in giustificazione, più che autentica spiegazione, del risultato di un processo algoritmico. Per rendere l'AI Act realmente efficace — e non semplice dichiarazione di principio — è necessario, dunque, un lavoro interpretativo coerente con gli obiettivi sostanziali del regolamento. In sua assenza, l'AI Act rischierebbe infatti di ridursi a un guscio formale, trasformandosi in un apparato burocratico di *compliance* che svuota di significato gli obiettivi stessi della norma. Un simile esito finirebbe per avvantaggiare soprattutto le grandi multinazionali, dotate delle risorse economiche e organizzative necessarie a sostenere gli oneri previsti, a svantaggio degli attori più piccoli<sup>29</sup>. Permane inoltre la questione dell'effettiva vincolatività del quadro: alla data odierna alcuni dei soggetti più influenti — ad esempio Meta — hanno già annunciato l'intenzione di non

26. Il tema dell'automazione dei processi di pubblica amministrazione è quantomai centrale e controverso, soprattutto se non supportato da investimenti in infrastrutture di IA di proprietà pubblica, dovendosi dunque affidare a sistemi prodotti da terzi. Su questi aspetti si veda Faini, Fernanda "Intelligenza artificiale, diritto e pubblica amministrazione", in A. D'Aloia (a cura di), *Intelligenza artificiale e diritto. Come regolare un mondo nuovo*. Franco Angeli, Milano

2020: 385-413. Cfr., anche, Ferrari, Mariangela "Il vantaggio della responsabilità concorsuale da uso"organizzato" di algoritmi" *Ragion pratica*. 2 (2021): 405-426,

27. Alcuni interessanti contributi recenti sul tema, sono Gintare Makauskaitė-Samuole. "Transparency in the Labyrinths of the EU AI Act: Smart or Disbalanced?." *Access to Justice in Eastern Europe* (2025), Holst, Luca; Lämmermann, Luis; Mayer, Valentin; Urbach, Nils; and Wendt, Domenik. "The Impact of the EU AI Act's Transparency Requirements on AI Innovation" (2024). *Wirtschaftsinformatik 2024 Proceedings*. 92.

28. Buttaboni, Carlotta and Floridi, Luciano, *A Taxonomy of AI Opacity in the EU: Rethinking Transparency, Traceability, Interpretability, and Explainability* (July 23, 2025). Available at SSRN: <https://ssrn.com/abstract=5364019> or <http://dx.doi.org/10.2139/ssrn.5364019>

29. Gintare Makauskaitė-Samuole, *op. cit.*

sottoscrivere il regolamento<sup>30</sup>. L'AI Act ambisce a mitigare i rischi sociali dell'intelligenza artificiale con riferimento ai diritti individuali; tuttavia, non sembra affrontare il cuore del problema, che è politico, economico e sistemico. Il nodo cruciale non risiede tanto nella tecnologia in sé — che può generare benefici significativi se prodotta e impiegata in modo responsabile e nell'interesse della collettività — quanto nella concentrazione delle infrastrutture digitali nelle mani di pochissimi poteri privati.

In sintesi, non basta dunque che l'algoritmo *possa* essere spiegato, occorre anche che vi sia un soggetto nelle condizioni normative ed epistemiche in grado di accedere alle risorse necessarie a detta spiegazione. Un aspetto cruciale dell'opacità algoritmica, specie nella sfera pubblica digitale, è quello che potremmo definire l'opacità ambientale o architetturale: non del singolo modello, ma dell'intero ecosistema algoritmico entro cui interagiamo quotidianamente (reti sociali, motori di ricerca, piattaforme di e-commerce). Le architetture algoritmiche di rete e piattaforma strutturano il flusso di informazioni, la visibilità dei contenuti e le modalità di interazione tra gli utenti. Esse costituiscono una sorta di infrastruttura non immediatamente percettibile che indirizza comportamenti e percezioni, spesso al di fuori della piena consapevolezza degli stessi utilizzatori. Allo stesso tempo, le scelte in fase di programmazione del codice e organizzazione delle logiche di piattaforma assumono una portata normativa iperpositiva, delimitando tramite regole tecniche il campo d'azione dei soggetti in rete. Per riprendere l'efficace espressione di Lawrence Lessig, negli ambienti digitali *code is law*<sup>31</sup>.

### 3 Bolle epistemiche e propaganda computazionale

L'opacità algoritmica assume dunque una dimensione critica nel contesto di un ambiente digitale che appare sempre più oligopolistico, in cui le infrastrutture per la raccolta, il volume e il valore dei dati prodotti — la nuova materia prima dell'economia digitale<sup>32</sup> — sono mani di pochissimi attori privati. Se difatti in merito a decisioni di carattere tecnico supervisionate da un decisore umano — ad esempio l'ambito medico o scientifico — appare praticabile una trasparenza che passi per robusti sistemi di *Explainable AI* e lo sviluppo di software proprietari ad applicazione specifica; il discorso si complica in riferimento ai sistemi di decisione non supervisionata<sup>33</sup>, di *General Purpose AI* e per quanto riguarda le architetture algoritmiche che regolano le grandi piattaforme della rete. Appare così egualmente urgente e difficoltoso adottare una linea politica e giuridica che faccia della limitazione dei grandi poteri privati terreno strategico di discussione e intervento. Ciò non solo per i ben noti effetti negativi che l'accentramento dei capitali ha sulla libertà e pluralità dei mercati, ma anche a causa dell'inedito squilibrio di potere epistemico e politico che tale dinamica provoca, disegnando uno spazio che — seppur ormai centrale nella vita individuale e collettiva della cittadinanza — resta sostanzialmente esterno a logiche e responsabilità di carattere pubblico.

Tale fenomeno solleva criticità anche per quanto riguarda l'autonomia cognitiva e decisionale dei soggetti in rete. Difatti, l'esplosione del fenomeno di Big Data ha inaugurato una fase radicalmente nuova nel campo della regolazione comportamentale "*soft*", caratterizzata da meccanismi di influenza che superano le tradizionali strategie di *nudging*<sup>34</sup>. Precedentemente esse potevano operare attraverso interventi statici e generali nell'architettura delle scelte, le pratiche odierne possono al contrario fare uso di sistemi algoritmici dinamici, auto-adattivi e costantemente aggiornati, capaci di modellare in tempo reale i contesti informativi degli individui. Questi sistemi sfruttano la capacità degli algoritmi di individuare correlazioni all'interno di ampi flussi di dati rendendo l'esperienza digitale altamente personalizzata e difficilmente decifrabile dagli utenti stessi.

30. "Meta won't sign EU's AI Code, but who will?" Euronews, 23 July 2025, <https://www.euronews.com/my-europe/2025/07/23/meta-wont-sign-eus-ai-code-but-who-will>.

31. Lessig, Lawrence. *Code and other laws of cyberspace*, New York, Basic Books, 1999.

32. Zuboff, Shoshana, *op. cit.*

33. Si pensi, ad esempio, alle rapidissime micro-decisioni che prende un veicolo a guida autonoma per minimizzare il rischio.

34. Il termine *nudge* letteralmente significa "colpetto" o "spintarella". Esso indica qualunque intervento sull'"architettura delle scelte" che orienta il comportamento in modo prevedibile senza vietare alternative né modificare in maniera rilevante gli incentivi economici. Rappresenta un approccio alla modifica comportamentale basato sulle scelte di design degli ambienti. In italiano si è affermata la traduzione "spinta gentile" proprio per sottolineare la natura non coercitiva dell'intervento, a differenza di leggi o divieti espliciti. L'elaborazione della teoria del *nudging* fa riferimento a Thaler, Richard H., and Cass R. Sunstein. *Nudge: Improving Decisions about Health, Wealth, and Happiness*. Yale University Press, 2008.

In tal modo, si afferma una forma di regolazione comportamentale che, pur mantenendo le apparenze di una spinta "morbida" e priva di coercizione formale, acquisisce un potere di condizionamento cognitivo inedito, opaco, particolarmente pervasivo e difficilmente limitabile dagli strumenti tradizionali<sup>35</sup>. Gli algoritmi di raccomandazione non si limitano a filtrare l'informazione, ma delimitano attivamente il campo epistemico degli utilizzatori finali. Sul piano strategico, la logica di questi sistemi resta solitamente ancorata alla massimizzazione dell'*engagement*, variabile che, nell'economia dell'attenzione, traduce direttamente il tempo speso in redditività pubblicitaria. Emblematico è il caso di TikTok: studi empirici dimostrano che l'algoritmo "For You" è in grado di inferire le preferenze di un nuovo utente nel giro di pochi minuti, orientando il 95 % dei contenuti mostrati sulla base di un modello predittivo che ottimizza il tempo di permanenza sulla piattaforma<sup>36</sup>. Audit indipendenti hanno rilevato che funzioni come follow e like esercitano un impatto decisivo sulla successiva selezione dei video, favorendo la formazione di *cluster* di interesse omogenei e quindi di potenziale isolamento e radicalizzazione in ambienti informativi co-prodotti<sup>37</sup>. Parallelamente, ricerche qualitativo-sperimentali evidenziano come molti giovani utenti percepiscano l'algoritmo quale entità insondabile e onnipotente, capace di decretare le fortune in rete di contenuti e utenti<sup>38</sup>. Non sorprende inoltre che studi di psicologia sociale segnalino come l'esposizione continuativa a feed personalizzati possa incidere sulla costruzione identitaria e sui valori interni degli utilizzatori<sup>39</sup>.

Fenomeni simili non sono limitati alle piattaforme social, ma attraversano in maniera ubiqua lo spazio digitale, pervadendo anche ambienti che, a uno sguardo superficiale, potrebbe sembrare intuitivamente esenti da siffatte dinamiche di massimizzazione dell'*engagement*. Nel campo *dei news recommender*, ad esempio, ciò si traduce in metriche che privilegiano il *click-through* e la permanenza a scapito di criteri editoriali di interesse pubblico, conducendo al serio rischio di restringere, nella pratica, la varietà e la qualità dell'informazione consumata dalla cittadinanza, comprimendo oltretutto il pluralismo delle fonti<sup>40</sup>.

Alla luce di siffatte dinamiche, l'opacità — e la correlata richiesta di trasparenza — si configura come snodo essenziale per affrontare sfide che incidono profondamente sulla sfera pubblica di Stati e cittadini. I fenomeni descritti conducono difatti a forme di rischio che non si limitano a quelle — ben documentate in letteratura — attinenti alla sfera dei diritti e delle libertà individuali<sup>41</sup>, ma che assumono natura sistemica, coinvolgendo i presupposti della democrazia e del diritto<sup>42</sup>. Anche in assenza di deliberate strategie manipolative, fenomeni

---

35. Karen Yeung descrive questo scarto qualitativo parlando di Hypernudge. "Hypernudging relies on highlighting algorithmically determined correlations between data items within data sets that would not otherwise be observable through human cognition alone (or even with standard computing support), thereby conferring 'salience' on the highlighted data patterns, operating through the technique of 'priming', dynamically configuring the user's informational choice context in ways intentionally designed to influence her decisions" Yeung, Karen, "Hypernudge". *Big Data as a mode of regulation by design, Information, Communication & Society*, 20:1 (2017), p. 122.

36. Zannettou, Savvas, et al. "Analyzing User Engagement with TikTok's Short Format Video Recommendations Using Data Donations." *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI '24)*, Association for Computing Machinery, 2024.

37. Boeker, Maximilian, and Aleksandra Urman. "An Empirical Investigation of Personalization Factors on TikTok." *Proceedings of the 16th ACM Conference on Web Science (WebSci '22)*, 2022: 159-167.

38. Felaco, Cristiano. "Making Sense of Algorithm: Exploring TikTok Users' Awareness of Content Recommendation and Moderation Algorithms." *International Journal of Communication* 19 (2025): 1081-1102.

39. Ionescu, Claudiu Gabriel, and Monica Licu. "Are TikTok Algorithms Influencing Users Self-Perceived Identities and Personal Values? A Mini Review." *Soc. Sci.*, 12. 8, 2023: 465.

40. Helberger, Natali. "On the Democratic Role of News Recommenders." *Digital Journalism* 7.8(2019): 993-1012

41. Negli ultimi anni il tema è stato al centro delle riflessioni critiche della letteratura, producendo un gran numero di contributi di valore. Fra di essi ricordiamo, per un inquadramento generale Bello, Barbara G. (In)giustizie digitali. *Un itinerario fra tecnologie e diritti*. Pisa, Pacini Giuridica, 2023; Sartor, Giovanni. "Artificial intelligence and human rights: between law and ethics." *Maastricht Journal of European and Comparative Law* 27.6 (2020): 705-719.

42. Durante, Massimo. *Potere computazionale*, Milano, Meltemi, 2019.

quali *echo chambers*<sup>43</sup> e *filter bubbles*<sup>44</sup> - vere e proprie bolle epistemiche<sup>45</sup> - contribuiscono a inquinare i processi di formazione dell'opinione, anche politica, giocando un ruolo significativo nella tendenza di radicalizzazione e frammentazione del discorso pubblico propria delle piattaforme digitali<sup>46</sup>. Parallelamente, il fenomeno della propaganda computazionale appare diffuso e quantomai difficoltoso da individuare e da riportare all'interno di una cornice legale e democratica. Con propaganda computazionale si intende "the use of algorithms, automation, and human curation to purposefully distribute misleading information over social media networks. These activities can feed into influence campaigns: coordinated, illegitimate efforts of a third state or non-state agent to affect democratic processes and political decision-making, including (but not limited to) election interference"<sup>47</sup>.

La propaganda computazionale rappresenta una rilevante evoluzione rispetto alle forme tradizionali di manipolazione dell'opinione pubblica, sia per dimensioni e pervasività sia per quanto riguarda le tecniche utilizzate. Diversamente dai modelli radiofonici e televisivi novecenteschi, che operavano perlopiù attraverso canali centralizzati e facilmente identificabili, la propaganda computazionale sfrutta l'architettura distribuita delle piattaforme digitali per creare quello che Woolley e Howard definiscono nei termini di ecosistema manipolativo<sup>48</sup>. La specificità di questa - relativamente nuova - modalità di comunicazione di massa risiede nella sua capacità di operare simultaneamente su tre livelli: la produzione automatizzata di contenuti attraverso algoritmi generativi (*political bot*), la distribuzione mirata attraverso sistemi di profilazione comportamentale (*microtargeting*)<sup>49</sup>, e l'amplificazione coordinata attraverso reti di account automatizzati che simulano consenso organico (*bot networks*). Profilazione e microtargeting rappresentano strumenti dall'incredibile efficacia, permettendo di confezionare contenuti estremamente accurati e personalizzati, accrescendo enormemente l'efficienza del messaggio rispetto a sistemi di propaganda tradizionali. La letteratura psicologica mostra infatti come l'esposizione a input mirati, calibrati sul profilo psico-comportamentale, sia in grado di aggirare la deliberazione riflessiva, generando reazioni pre-razionali che sfuggono al controllo conscio<sup>50</sup>. (Ienca, *On Artificial Intelligence and Manipulation*, Topoi, 2023, p. 835)

#### 4 Agentività e sovranità epistemica dinnanzi alla sfida dell'opacità

Si è dunque visto come, allo stato attuale, ambienti digitali retti da architetture algoritmiche opache siano effettivamente in grado di produrre, anche involontariamente, "semplicemente" aderendo alle logiche di massimizzazione del profitto della piattaforma una distorsione epistemica dei soggetti in rete. In aggiunta, tali spazi si prestano ad essere usati quale veicolo di deliberata manipolazione politica. La pervasività di siffatte tecnologie fa emergere in maniera concreta le ingerenze da parte di terzi (stati o privati) all'interno dei processi democratici. Ciò non sarebbe ovviamente un fenomeno nuovo, ma le caratteristiche intrinseche delle

43. Si parla di *echo chamber* quando una rete di relazioni omofile e di interazioni ripetute induce la circolazione quasi esclusiva di contenuti conformi alle convinzioni condivise dal gruppo, con progressiva esclusione (o delegittimazione) di fonti alternative. In tale contesto, opinioni già affini vengono continuamente "rispecchiate" e rafforzate, producendo coesione interna e, spesso, polarizzazione verso l'esterno.

44. Il termine *filter bubble* designa la sfera informativa personalizzata generata dagli algoritmi di raccomandazione e di ranking: sulla base dei profili di comportamento, il sistema seleziona e ordina contenuti "pertinenti", occultandone altri potenzialmente rilevanti ma incongruenti con le preferenze inferite. Connesso è anche il fenomeno dei cosiddetti *rabbit hole*, percorsi di raccomandazione che spingono progressivamente verso contenuti sempre più estremi.

45. Coeckelbergh, Mark. Democracy, epistemic agency, and AI: political epistemology in times of artificial intelligence." *AI and Ethics* 3 (2023): 1341-1350. Su questi temi fondamentale citare Praiser, Eli. *The Filter Bubble: What the Internet Is Hiding From You*. Londra, Penguin Books, Londra, 2011.

46. Rilevantissimo su questi temi, nel contesto del discorso pubblico attorno alla pandemia COVID-19, il contributo di Schmidt, Ana L., et. Al. "Polarization of the vaccination debate on Facebook" *Vaccine* 36.25 (2018): 3606-3612.

47. Bentzen N. (2018), *Computational Propaganda Techniques*, European Parliamentary Research Service – At a Glance.

48. Woolley, Samuel C., and Philip N. Howard, a cura di. *Computational Propaganda: Political Parties, Politicians, and Political Manipulation on Social Media*. Oxford University Press, 2019

49. Lagioia, Francesco e Sartor, Giovanni. "Profilazione e decisione algoritmica: dal mercato alla sfera pubblica". *Federalismi. Rivista di diritto pubblico italiano, comparato, europeo*. 11(2020): 103-104. Cfr anche Coniglione, Casimiro. "L'utilizzo dei big data in ambito politico-elettorale e il loro impatto sulla democrazia rappresentativa". *Nomos. Le attualità del diritto*. 1(2023).

50. Ienca, *On Artificial Intelligence and Manipulation*, Topoi, 2023, p. 835

tecnologie che veicolano attualmente l'informazione rendono la scala di tali processi potenzialmente senza precedenti.

In questo contesto l'opacità assume le caratteristiche di dispositivo volto a proteggere una delle principali categorie legittimanti del potere epistemico assunto dalle piattaforme: quella della – presunta – neutralità e oggettività della tecnologia da esse adoperata.<sup>51</sup> L'opacità algoritmica impedisce infatti di visualizzare con certezza le logiche interne dei sistemi automatici, occultandone le scelte valoriali e architettoniche, permettendo un approccio analitico unicamente *ex post*, a partire dagli effetti prodotti dai sistemi, e non *ex ante*, a partire dalla struttura interna degli stessi.

La situazione descritta delinea un orizzonte in cui a essere messo a rischio è uno dei presupposti fondamentali di una democrazia in salute: l'*epistemic agency* degli individui che la compongono. Quest'ultima, seguendo la definizione data da Coeckelbergh, concerne "the question regarding control over one's beliefs and how these beliefs are formed and revised"<sup>52</sup>.

Tale premessa costituisce il fondamento cognitivo della vita democratica. Per il cittadino, poter riflettere sulle ragioni che sorreggono le proprie opinioni politiche e sottoporle a discussione pubblica è condizione necessaria per processi deliberativi critici e consapevoli. Essi richiedono difatti soggetti in grado di padroneggiare il proprio sapere politico, in altre parole "political democratic agency seems to rely on epistemic agency"<sup>53</sup>. Quest'esigenza di agentività epistemica non è dunque unicamente di carattere individuale, ma, al contrario, riguarda la collettività quale soggetto politico, declinandosi dunque nei termini di sovranità epistemica. Quest'ultima è intesa nei termini dell'abilità "of a group (...) to exert control and autonomy over their own epistemic processes without being subject to external forms of domination or oppression"<sup>54</sup>. Si configura dunque come la capacità dei soggetti politici di costituirsi come agenti attivi nei processi di conoscenza, sottraendosi alle logiche eterodirette prefigurate imposte dai dispositivi di governabilità; la facoltà di una collettività di negoziare pubblicamente i criteri di formazione della verità socialmente operativa e di contestarne gli esiti. Quando tali criteri si inscrivono in architetture opache e private la sovranità epistemica tende a allontanarsi dal contesto pubblico e democratico. Essa, tuttavia, non può essere pensata come proprietà naturale, ma deve essere intesa come progettualità, richiedendo specifiche condizioni sociali, materiali e culturali per potersi dispiegare<sup>55</sup>. La prospettiva della sovranità epistemica non può quindi limitarsi alla rivendicazione di forme astratte di partecipazione ai processi conoscitivi, ma deve necessariamente confrontarsi con la questione delle condizioni materiali e tecnologiche che soggiacciono alla produzione dei saperi. Come sottolinea Oliveira Mendes Pinto, "epistemic sovereignty emphasizes the importance of control, access, and autonomy over the production of information and knowledge and challenges power structures and dominant epistemologies"<sup>56</sup>.

## 5 Conclusioni

L'analisi condotta ha cercato di evidenziare in quali termini l'opacità algoritmica nel contesto del capitalismo delle piattaforme rappresenti una sfida sistemica ai fondamenti epistemici della democrazia contemporanea. La convergenza tra le logiche di profitto delle piattaforme digitali e architetture algoritmiche opache ha gen-

51. Al contrario, la prospettiva da adoperare, è quella della co-produzione di scienza e società, sulla scia dei risultati degli Science and Technology Studies. Sul concetto e sui modi della "co-produzione" si veda S. Jasanoff (a cura di), *States of Knowledge: The Co-Production of Science and Social Order*, Routledge, London-New York, 2004. Altri testi utili legati al tema sono T. Bonazzi, *Tecnologia e democrazia. Saggio sulla razionalità moderna*, Il Mulino, Bologna, 1995; M. Callon, *Socio-technical networks and irreversibility*, in J. Law (ed. by), *A Sociology of Monsters: Essays on Power, Technology and Domination*, Routledge, London, 1991, pp. 132-161; R. Finelli, *La società e la scienza. Per una teoria critica della tecnoscienza*, DeriveApprodi, Roma, 2010; B. Latour, *La scienza in azione. Introduzione alla sociologia della scienza*, (1987), Edizioni di Comunità, Milano, 2000.

52. Coeckelbergh, Mark, *op. cit.*, p. 1342.

53. Ivi, p. 1343.

54. Oliveira, Thaiane, e Tatiane Mendes Pinto. "Knowledge and emancipation: From epistemic injustice to digital and epistemic sovereignty in Latin America." *Journal of Digital Social Research* 6.3(2024): 40–58. In particolare p.43

55. Medina, *The Epistemology of Resistance*, Oxford University Press, 2013, p. 84.

56. Oliveira e Mendes Pinto, *op. cit.*, p. 53

erato un ecosistema informativo che compromette strutturalmente l'agentività degli individui, minacciando di trasformare soggetti epistemici attivi in oggetti di *nudging* continuo e propaganda computazionale. La scala e la pervasività di tali processi, amplificate dalle caratteristiche intrinseche delle tecnologie digitali, conferiscono a questi fenomeni una portata potenzialmente senza precedenti. La distinzione tra opacità strutturale e opacità socialmente costituita illustra come il problema non sia riducibile a questioni meramente tecniche di interpretabilità degli algoritmi, ma investa le modalità stesse attraverso cui l'informazione viene prodotta, diffusa e fruita negli spazi digitali. L'opacità si configura infatti come dispositivo strategico che consente alle piattaforme di mantenere in piedi la narrativa della neutralità tecnologica, occultando le scelte valoriali e gli interessi particolari che orientano i sistemi algoritmici. Di fronte a tali processi, la rivendicazione di una sovranità epistemica si presenta come requisito per la preservazione dei presupposti democratici. Il percorso che vi conduce non può limitarsi a interventi normativi di tipo reattivo, ma richiede un ripensamento proattivo dei modelli di regolamentazione dell'innovazione tecnologica che tenga in serio conto la tutela dell'autonomia cognitiva dei cittadini. Se un'educazione critica e consapevole al digitale emerge come prerequisito fondamentale per l'esercizio di un'autonomia sia cognitiva sia politica, essa non può essere considerata autonoma e sufficiente, se non accompagnata da trasformazioni strutturali di indirizzo economico e politico che affrontino le asimmetrie di potere e conoscenza generate – e presupposte – dal capitalismo delle piattaforme.

## 6 Bibliografia

Ali, Sajid, et al. "Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence." *Information Fusion* 99 (2023): 101805.

Amato, Sabrina. *Biodiritto 4.0. Intelligenza artificiale e nuove tecnologie*. Giappichelli, Torino, 2020.

Bello, Barbara G. *(In)giustizie digitali. Un itinerario fra tecnologie e diritti*. Pacini Giuridica, Pisa, 2023.

Bentzen, Naja. *Computational Propaganda Techniques*. European Parliamentary Research Service – At a Glance, 2018.

Boge, Florian J. "Two Dimensions of Opacity and the Deep Learning Predicament." *Minds and Machines* 32.1(2022): 43-75.

Boeker, Maximilian, and Aleksandra Urman. "An Empirical Investigation of Personalization Factors on TikTok." *Proceedings of the 16th ACM Conference on Web Science (WebSci '22)*, 2022: 159-167.

Boden, Margaret A. "GOFAI." *The Cambridge Handbook of Artificial Intelligence*. Keith Frankish e William M. Ramsey (a cura di). Cambridge University Press, Cambridge, 2014, 89–107.

Bonazzi, Tommaso. *Tecnologia e democrazia. Saggio sulla razionalità moderna*. Il Mulino, 1995.

Brighi, Raffaella e Zullo, Silvia. *Filosofia del diritto e nuove tecnologie. Prospettive di ricerca tra teoria e pratica*. Aracne, Roma, 2015.

Burrell, Jane. "How the machine 'thinks': Understanding opacity in machine learning algorithms." *Big Data & Society* 3.1 (2016).

Buttaboni, Carlotta e Floridi, Luciano. "A Taxonomy of AI Opacity in the EU: Rethinking Transparency, Traceability, Interpretability, and Explainability." *SSRN*, 23 July 2025.

Casadei, Thomas e Stefano Pietropaoli. "Intelligenza artificiale: l'ultima sfida per il diritto?" *Diritto e tecnologie informatiche. Questioni di informatica giuridica, prospettive istituzionali e sfide sociali*. Thomas Casadei e Stefano Pietropaoli (a cura di). Wolters Kluwer, 2024, 259–274.

Casadei, Thomas; Marzocco Valeria e Zullo Silvia. *La didattica del diritto. Metodi, strumenti e prospettive* (seconda edizione). Pancini Editore, Pisa, 2021

Christin, Angèle. "The ethnographer and the algorithm: beyond the black box." *Theory and Society*. 49(2020): 897-918.

- Coeckelbergh, Mark. Democracy, epistemic agency, and AI: political epistemology in times of artificial intelligence." *AI and Ethics* 3(2023): 1341-1350cfr. nota per pp. 1342-1343.
- Coniglione, Casimiro. "L'utilizzo dei big data in ambito politico-elettorale e il loro impatto sulla democrazia rappresentativa". *Nomos. Le attualità del diritto*. 1(2023).
- Durante, Massimo. *Potere computazionale. L'impatto delle ICT su diritto, società, sapere*. Meltemi, Milano, 2019.
- Facchini, Alessandro, and Alberto Termine. "Towards a Taxonomy for the Opacity of AI Systems." *Philosophy and Theory of Artificial Intelligence*. 63(2022).
- Faini, Fernanda "Intelligenza artificiale, diritto e pubblica amministrazione", in A. D'Aloia (a cura di), *Intelligenza artificiale e diritto. Come regolare un mondo nuovo*. Franco Angeli, Milano, 2020: 385-413.
- Felaco, Cristiano. "Making Sense of Algorithm: Exploring TikTok Users' Awareness of Content Recommendation and Moderation Algorithms." *International Journal of Communication*. 19(2025): 1081-1102.
- Felzmann, Helke; Fosch-Villaronga Eduard; Lutz, Christoph e Tamò-Larrieux, Aurelia."Towards Transparency by Design for Artificial Intelligence." *Science and Engineering Ethics*. 26(2020): 3333 - 3361.
- Ferrari, Mariangela "Il vantaggio della responsabilità concorsuale da uso"organizzato" di algoritmi". *Ragion pratica*. 2(2021): 405-426
- Finelli, Roberto. *La società e la scienza. Per una teoria critica della tecnoscienza*. DeriveApprodi, 2010.
- Fioriglio, Gianluigi. *Informatica medica e diritto. Un'introduzione*. Mucchi, Modena, 2020
- Fioriglio, Gianluigi. "La Società algoritmica fra opacità e spiegabilità: profili informatico-giuridici." *Ars interpretandi. Rivista di ermeneutica giuridica*. 1(2021): 53-67.
- Floridi, Luciano. *La quarta rivoluzione industriale. Come l'infosfera sta trasformando il mondo*. Cortina, Milano, 2017.
- Floridi, Luciano. *Etica dell'intelligenza artificiale: sviluppi, opportunità, sfide*. Cortina, Milano, 2022.
- Gintare Makauskaite-Samuole, "Transparency in the Labyrinths of the EU AI Act: Smart or Disbalanced?" *Access to Justice in Eastern Europe*. 8.2(2025): 1-31.
- Helberger, Natali. "On the Democratic Role of News Recommenders." *Digital Journalism* 7.8(2019): 993-1012.
- Holst, Luca; Lämmermann, Luis; Mayer, Valentin; Urbach, Nils; and Wendt, Domenik. "The Impact of the EU AI Act's Transparency Requirements on AI Innovation." *Wirtschaftsinformatik 2024 Proceedings*. 92(2024).
- Humphreys, Paul. "The Philosophical Novelty of Computer Simulation Methods." *Synthese* 169.3(2009): 615-626.
- Ienca, Carlo. "On Artificial Intelligence and Manipulation." *Topoi* 42.3(2023): 833-842
- Ionescu, Claudiu G. e Monica Licu. "Are TikTok Algorithms Influencing Users' Self-Perceived Identities and Personal Values? A Mini Review." *Soc. Sci.* 12.8(2023): 465.
- Jasanoff, Sheila (a cura di). *States of Knowledge: The Co-Production of Science and Social Order*. Routledge, Londra, 2006.
- Lagioia, Francesca e Sartor, Giovanni. "Profilazione e decisione algoritmica: dal mercato alla sfera pubblica". *Federalismi. Rivista di diritto pubblico italiano, comparato, europeo*. 11(2020): 85-110.
- Latour, Bruno. *La scienza in azione. Introduzione alla sociologia della scienza*. Edizioni di Comunità, Roma 2000.
- Lessig, Lawrence. *Code and other laws of cyberspace*. Basic Books, New York, 1999.

- Loi, Michele, et al. "Transparency as design publicity: explaining and justifying inscrutable algorithms." *Ethics and Information Technology*. 23.3(2019): 253–263.
- Mann, Sara, et al. "Sources of Opacity in Computer Systems: Towards a Comprehensive Taxonomy." *IEEE 31st International Requirements Engineering Conference Workshops (REW)*, 2023: 334–344.
- Martoni, Michele. "Un'autonomia ostacolata: Limiti cognitivi, incompetenze e design ingannevoli nella trasformazione digitale." *Sociologia del diritto* 51.1(2024): 7–32.
- Medina, José. *The Epistemology of Resistance*. Oxford University Press, New York, 2013.
- Meta won't sign EU's AI Code, but who will?* Euronews, 23 July 2025.
- Oliveira, Thaiané, and Tatiane Mendes Pinto. "Knowledge and emancipation: From epistemic injustice to digital and epistemic sovereignty in Latin America." *Journal of Digital Social Research* 6.3(2024): 40–58.
- Oliveri, Federico. *Machina Mundi. Per una regolazione democratica dei poteri digitali*. Mucchi editore, Modena, 2025.
- Palazzani, Laura. *Tecnologie dell'informazione e intelligenza artificiale. Sfide etiche al diritto*. Studium, Roma 2020.
- Pascuzzi, Giovanni. *La cittadinanza digitale. Competenze, diritti e regole per vivere in rete*. Il Mulino, Bologna, 2021
- Pasquale, Frank. *The Black Box Society: The Secret Algorithms That Control Money and Information*. Harvard University Press, Cambridge/Londra, 2015.
- Pietropaoli, Stefano. "Verso un legislatore non umano? Brevi riflessioni su alcuni problemi di diritto computazionale." *Osservatorio sulle fonti* 2(2022): 397-410.
- Praiser, Eli. *The Filter Bubble: What the Internet Is Hiding From You*. Londra, Penguin Books, Londra, 2011.
- Sajid, Ali; Abuhmedb, Tamer; El-Sappaghb, Shaker; Muhammadc, Khan; Alonso-Morald, José M.; Confalonierie, Roberto; Guidottif, Riccardo; Del Serg, Javier; Díaz-Rodríguez, Natalia e Herrera, Francisco. "Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence." *Information Fusion* 99(2023): 101805.
- San Pedro García, Inaki. "Degrees of Epistemic Opacity." *Teorema. International Journal of Philosophy* 43.2(2024): 5-21.
- Sapienza, Salvatore. *Decisioni algoritmiche e diritto*. Giuffrè Francis Lefebvre, Milano, 2024.
- Sartor, Giovanni. *L'intelligenza artificiale e il diritto*. Giappichelli, Torino 2022.
- Sartor, Giovanni. "Artificial intelligence and human rights: between law and ethics." *Maastricht Journal of European and Comparative Law* 27.6(2020): 705-719.
- Scagliarini, Simone. "I diritti costituzionali nell'era di internet: cittadinanza digitale, accesso alla rete e net neutrality". In *Diritto e tecnologie informatiche. Questioni di informatica giuridica, prospettive istituzionali e sfide sociali*. Thomas Casadei e Stefano Pietropaoli (a cura di). Wolters Kluwer, Alphen aan den Rijn, 2022: 3 -13.
- Schmidt, Ana Lucía Schmidt; Zollo, Fabiana; Scala, Antonio; Betsch Cornelia; Quattrocchi, Walter. "Polarization of the vaccination debate on Facebook" *Vaccine* 36.25(2018): 3606-3612.
- Schmidt, Eva et al. "The Epistemic Cost of Opacity: How the Use of Artificial Intelligence Undermines the Knowledge of Medical Doctors in High-Stakes Contexts". *Philos. Technol.* 38.5(2025).
- Thaler, Richard H., and Cass R. Sunstein. *Nudge: Improving Decisions about Health, Wealth, and Happiness*. Yale University Press, New Haven 2008.

Woolley, Samuel C., and Philip N. Howard, eds. *Computational Propaganda: Political Parties, Politicians, and Political Manipulation on Social Media*. Oxford University Press, New York, 2019.

Yeung, Karen. "Hypernudge. Big Data as a mode of regulation by design." *Information, Communication & Society* 20.1(2017): 118-136.

Zannettou, Savvas; Nemes-Nemeth, Olivia; Ayalon, Oshrat; Goetzen, Angelica; Gummadi, Krishna P.; Redmiles, Elissa M. e Roesner, Franziska. "Analyzing User Engagement with TikTok's Short Format Video Recommendations Using Data Donations." *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI '24)*, Association for Computing Machinery, 2024: 1-16.

Zuboff, Shoshana. *Il capitalismo della sorveglianza. Il futuro dell'umanità nell'era dei nuovi poteri*. LUISS University Press, 2020.

Winter, Peter D. e Carusi Annamaria "(De)troubling transparency: artificial intelligence (AI) for clinical applications" *Medical Humanities* 49(2023):17-26.