



Editoriale

doi: 10.82015/NNR.2025.100112

Genere, potere e rischio nell'era dell'Intelligenza artificiale

Luigi Di Cataldo*

1. L'Intelligenza Artificiale come costrutto socio-tecnico: oltre il mito della neutralità algoritmica

L'Intelligenza Artificiale (IA) ha inaugurato una trasformazione strutturale che permea in profondità ogni aspetto della vita sociale, economica e politica, ridefinendo in modo radicale i paradigmi della convivenza civile e dell'allocazione delle risorse. Per lungo tempo, la narrazione dominante ha dipinto l'innovazione tecnologica e algoritmica come una forza intrinsecamente neutrale, oggettiva e infallibile, capace di superare i limiti cognitivi umani in un trionfo di pura razionalità computazionale. Tuttavia, la ricerca scientifica degli ultimi anni ha progressivamente decostruito questo "mito della neutralità", rivelando la vera natura delle tecnologie computazionali recenti che consistono in costrutti socio-tecnici complessi, opachi e densi di implicazioni valoriali (Winner, 1980; Kitchin, 2017). In un classico imprescindibile degli studi sul tema, Frank Pasquale (2015) propone il concetto di "Black Box Society" in cui algoritmi imperscrutabili prendono decisioni vitali sui cittadini, celando dietro il segreto industriale le proprie logiche.

* Vicedirettore di Neural Nexus Review di E.N.I.A. (Ente Nazionale per l'Intelligenza Artificiale); Professore a contratto, CIAPI & Dipartimento di Scienze Politiche e Sociali, Università di Catania; Post-doc researcher, Dipartimento di Scienze Sociali e Politiche, Università di Milano; Componente del comitato scientifico-legale di Fondazione Ente Nazionale per l'Intelligenza Artificiale (ENIA); Consigliere nazionale della Società Italiana del Dottorato di Ricerca.

Orcid: <https://orcid.org/0000-0002-4304-5562> - E-mail: luigi.dicataldo@unimi.it.



Entro questo scenario, emergono forme di potere inedite e sganciate da ogni forma di responsabilità giuridica, sociale e democratica (Campolo & Crawford, 2020), che impongono la necessità di ripensare i processi di sviluppo tecnologico attraverso lenti partecipative. Gli algoritmi e le intelligenze artificiali, infatti, leggono la realtà, la riflettono e la filtrano, riproducendo e amplificando le dinamiche sociali e le gerarchie preesistenti (O'Neil, 2016).

Apprendendo da enormi set di dati che contengono le tracce storiche delle nostre disuguaglianze, l'IA rischia di automatizzare le discriminazioni se non governata proattivamente attraverso il coinvolgimento diretto della società civile.

Il trinomio "genere, potere e rischio" continua a rappresentare una chiave euristica rilevante per analizzare la trasformazione tecnologica in corso, comprendere criticamente e affrontare le sfide del nostro tempo. Neural Nexus Review offre questo numero monografico per analizzare l'intersezione tra la variabile di genere e le architetture del potere digitale, contribuendo a svelare in che modo tali dinamiche generino nuovi rischi sistemici e richiedano un'inedita e partecipata governance istituzionale (Crawford, 2021). Al contempo, il presente numero ospita tre *position paper* della Fondazione Ente Nazionale per l'Intelligenza Artificiale che illustrano alcuni framework metodologici inediti con il proposito di sostenere la governance etica dell'IA, mettendo a disposizione di policy-maker, professionisti legali, imprese e ricercatori strumenti pragmatici per ancorare a terra i concetti accademici e guidare uno sviluppo tecnologico che sia operativamente equo, responsabile e democraticamente sostenibile.

2. Asimmetrie algoritmiche: ridefinire il potere, calcolare il rischio, svelare i pregiudizi

L'approfondimento delle questioni che legano le categorie del genere e del rischio alla questione politica del potere nella società contemporanea implica, in primo luogo,



l'immergersi nei meccanismi invisibili che guidano le decisioni automatizzate e presuppone una rigorosa e aggiornata definizione di quest'ultimo.

Come teorizzato da Maas (2022), il potere digitale si concretizza in una dinamica relazionale, profondamente asimmetrica, mediata da tecnologie computazionali che rendono gli utenti "dipendenti", in modo effettivo e sistematico, dagli sviluppatori e dalle architetture tecniche per soddisfare i propri obiettivi quotidiani, scivolando in una condizione di marcata vulnerabilità e subalternità strutturale. Il potere digitale si cristallizza nel controllo esclusivo su risorse ed eventi (Kuźmicz, 2025), evolvendo in quello che De Gregorio (2023) definisce il "potere normativo dell'intelligenza artificiale": la capacità di pochi attori privati e transnazionali di dettare regole, standardizzare comportamenti e sostituirsi al decisore pubblico attraverso deleghe tecnologiche di cui sfugge il controllo democratico.

In questa dinamica di dominio, il rischio subisce una mutazione ontologica. Recuperando la celebre lezione di Ulrich Beck, esposta nel volume "La società del rischio" (1992), che presenta i pericoli moderni come forme sistemiche di vulnerabilità fabbricate dallo sviluppo tecno-scientifico. L'egemonia economica delle big tech, che controllano le piattaforme digitali, i big data e le intelligenze artificiali, composte da team di sviluppo debolmente rappresentativi della diversità sociale e professionalmente poco socializzati alle implicazioni sociali, sono ormai in grado di dettare regole, influenzare i comportamenti e definire standard morali, sociali ed economici. L'automazione decisionale in situazioni di potere strutturalmente asimmetriche trasforma la portata del rischio da individuale, quindi associato a un processo decisionale "unico" secondo le proprie logiche giustificatrici, a sistemico, in quanto associato a processi decisionali che si ripetono in modo automatico, imperscrutabili nelle proprie dinamiche interne e che sono in grado di colpire chiunque in ogni luogo e in qualsiasi momento.

Numerosi studi empirici dimostrano come i sistemi di Machine Learning e Natural



Language Processing assorbono sistematicamente bias patriarcali (Bolukbasi et al., 2016). Ne sono prove eclatanti i tassi di errore inaccettabili nei software di riconoscimento facciale per le donne afrodiscendenti (Buolamwini & Gebru, 2018), oppure anche le penalizzazioni che subiscono i profili femminili da parte dei sistemi di recruiting che integrano soluzioni di IA (Dastin, 2018; West et al., 2019). Queste distorsioni non sono fallimenti tecnici, ma la perfetta codificazione di un secolare squilibrio di potere che non può che essere scardinato tramite la democratizzazione e la partecipazione nelle attività e fasi di progettazione algoritmica.

La delega tecnologica nei processi decisionali (dal welfare alla giustizia penale, passando per il mondo del lavoro fino ai servizi bancari e assicurativi) richiede, dunque, una robusta strategia di gestione del rischio.

Il Regolamento (UE) 2024/1689 del Parlamento europeo e del Consiglio, del 13 giugno 2024, che stabilisce regole armonizzate sull'intelligenza artificiale (c.d. AI Act), adotta proprio quest'accezione, implementando un quadro modulato di gestione del rischio per proteggere i diritti fondamentali, lo Stato di diritto e le istituzioni democratiche (Schuett, 2024; Cheatham et al., 2019). Ciò nonostante, resta ancora vivo e fortemente avvertito il bisogno di evidenze empiriche, quadri interpretativi e strumenti concreti in grado di supportare efficacemente la governace dell'IA.

3. Oltre l'orizzonte attuale: tre macro-direttrici per le ricerche future

La crescente complessità di queste sfide sollecita una riflessione che sappia andare oltre la mera diagnosi del sintomo, tracciando direttrici di indagine adatte ad accompagnarci verso la costruzione di strumenti di governance dinamici e che siano al contempo efficaci.

La prima, e fondamentale, macro-area coinvolge le questioni che riguardano la partecipazione in fase di progettazione, implementazione e valutazione nell'ottica di



realizzare un futuro socio-digitale equo e condiviso. Gli studi futuri dovranno superare gli approcci meramente tecnici e quantitativi all'equità algoritmica, la cui limitatezza di fronte alle emergenti asimmetrie di potere è stata ampiamente documentata. Come evidenziato dai recenti studi di Gibin e colleghi (2025), occorre costruire un modello sociotecnico "laico" per negoziare nuovamente le relazioni sociali, coinvolgendo attivamente i non esperti, gli utenti finali e le comunità storicamente marginalizzate. La partecipazione diretta nella progettazione – dalla concettualizzazione all'etichettatura dei dati – non è un mero esercizio di stile, ma un'indicazione di metodo capace di decostruire il pregiudizio intersezionale e restituire potere a chi subisce l'impatto di tali tecnologie (D'Ignazio e Klein, 2020; Benjamin, 2019; Noble, 2018).

Nella seconda macro-area rientrano gli orizzonti della governance democratica e del costituzionalismo digitale. L'espansione tecnologica dei poteri privati solleva interrogativi cruciali sulla tenuta delle nostre democrazie, che coinvolgono il rapporto fra i poteri dello Stato (Michaels, 2020), i rapporti fra governanti e governati, nonché i rapporti fra privati (Kuźmicz, 2025). Interrogativi che potranno essere affrontati anche attraverso la partecipazione multi-stakeholder e il monitoraggio civico (De Gregorio, 2023).

La terza macro-area si focalizza su progettazione e operatività tecnologica. La ricerca dovrà condurci verso l'istituzionalizzazione di modalità progettuali e operative che ribaltino la logica top-down, passando dall'illusione di una tecnologia progettata "per" le persone a una tecnologia costruita "con" e "dalle" persone (Costanza-Chock, 2020; Wajcman, 2004). Questa direttrice dovrà indagare come integrare stabilmente le voci dei soggetti marginalizzati, disinnescando alla radice il rischio che l'innovazione si trasformi in uno strumento di oppressione e/o di dipendenza irreversibile (Yudkowsky, 2008; Müller, 2016).



4. Il contributo di questo numero: una mappa essenziale per orientarsi nel futuro digitale

I lavori ospitati in questo numero, selezionati per il loro impatto teorico e metodologico, si inseriscono in questo vivace dibattito multidisciplinare, fornendo ai lettori preziosi strumenti per comprendere e analizzare la transizione intelligente (Di Cataldo, 2025).

Aprire la raccolta l'articolo *«Dinamiche di potere, rapporti di genere e gestione del rischio nell'era dell'Intelligenza artificiale»* di Michele Di Salvo, che esplora l'impatto dell'IA sulle disuguaglianze di genere, spiegando come i pregiudizi vengano codificati negli algoritmi attraverso dataset storicamente viziati e sottolineando l'esistenza di una cultura industriale tech fortemente maschilista. L'Autore, inoltre, esamina i rischi legati alla sorveglianza biometrica dei corpi femminili, all'uso dell'IA nel welfare e alla riproduzione di stereotipi nei modelli linguistici e generativi. Per contrastare questa deriva, l'Autore propone lo sviluppo di un'etica femminista e radicalmente intersezionale, fondata su pratiche di design partecipativo e su una governance democratica capace di scardinare le dinamiche di dominio.

Segue il saggio *«Asimmetrie di potere emergenti. Una prospettiva europea per il futuro digitale»* di Roberto Magnani, che indaga le profonde ripercussioni dell'IA sulle dinamiche di potere contemporanee, esplorando l'evoluzione della cittadinanza in una dimensione sempre più computazionale e valutandone le relative sfide etiche, politiche e sociali. L'Autore propone una rilettura critica delle odierne strategie tecnologiche internazionali, che mette in guardia dal rischio di una concentrazione oligopolistica del potere digitale e dalla transizione da una semplice "economia dei dati" a una più pervasiva "economia delle intenzioni". L'Autore, infine, definisce la strategia dell'Unione europea come un modello sistemico alternativo, fondato sui principi di trasparenza, sovranità digitale e pluralismo culturale, sostenendo l'urgenza di un processo di alfabetizzazione in quanto presupposto irrinunciabile per una governance



autenticamente democratica.

Il terzo contributo, *«L'intelligenza artificiale è maschio o femmina?»* di Federica De Stefani, svela le dinamiche di potere e di genere insite nella progettazione, nel linguaggio e nell'utilizzo quotidiano delle tecnologie intelligenti. L'analisi si snoda lungo due direttrici principali: (i) esplora il fenomeno dell'antropomorfizzazione degli assistenti virtuali, mettendo in evidenza come le interfacce conversazionali interiorizzino e proiettino specifici stereotipi di genere, normalizzando ruoli subalterni; (ii) discute la questione della distorsione nei set di addestramento, che riproduce i pregiudizi storici e le asimmetrie sociali pre-esistenti. L'Autrice dimostra in modo inequivocabile come l'IA non sia uno strumento neutrale e che, al contrario, tenda piuttosto a cristallizzare e persino amplificare le gerarchie culturali preesistenti, e invoca un ripensamento strutturale e radicale dei paradigmi cognitivi e tecnici attuali così da abbracciare un orizzonte di sviluppo pienamente equo e plurale.

Nella sezione *«I position paper della Fondazione ENIA»*, che chiude questo numero di Neural Nexus Review dedicato al tema *«Genere, potere e rischio nell'era dell'intelligenza artificiale»*, sono raccolti tre documenti strategici e programmatici formalmente elaborati e promossi dalla Fondazione Enia. I documenti esprimono la visione ufficiale della Fondazione e le raccomandazioni corrispondenti in merito alle tematiche oggetto di trattazione, configurandosi come elementi di raccordo tra la riflessione teorica, l'analisi tecnica e l'azione istituzionale.

La Sezione si apre con il position paper *«Governare l'intelligenza artificiale: il Modello RICE»* firmato da Stefano Gorla (co-direttore del Dipartimento Etica dell'AI di Fondazione ENIA), Valeria Lazzaroli (Presidente di Fondazione Enia) e Luigi Di Cataldo (Vice-direttore di Neural Nexus Review e componente del comitato scientifico-legale di Fondazione Enia). Il Modello RICE rappresenta un framework metodologico applicato che offre a policy-maker, giuristi e società civile una bussola pragmatica per ancorare a terra i concetti accademici e guidare uno sviluppo tecnologico che sia operativamente



equo, responsabile e democraticamente sostenibile.

La Sezione prosegue con il position paper *«Un confronto tra standard ISO dedicati all'Etica dell'IA»* di Stefano Gorla (co-direttore del Dipartimento Etica dell'AI di Fondazione ENIA) e Valeria Lazzaroli (Presidente di Fondazione Enia). Gli autori teorizzano una sinergia indispensabile tra la regolamentazione cogente, che fissa i confini legali e le classificazioni di rischio, e gli standard ISO (ISO 26000; ISO/IEC 24368; ISO/IEC 42001), che forniscono le infrastrutture organizzative per tradurre la legge in prassi.

La Sezione si chiude con il position paper *«Autovalutazione etica dei sistemi di intelligenza artificiale. Un framework di governance, rischio e responsabilità organizzativa»* di Arianna Ciracò (co-direttrice del Dipartimento Etica dell'AI di Fondazione ENIA), Stefano Gorla (co-direttore del Dipartimento Etica dell'AI di Fondazione ENIA), Valeria Lazzaroli (Presidente di Fondazione ENIA) e Rebecca Paternostro (componente del Comitato etico di Fondazione ENIA). Il documento propone un framework operativo di autovalutazione etica per tradurre i principi normativi e le linee guida UNESCO in prassi aziendali misurabili. Attraverso un sistema analitico ponderato su 19 classi di principi, il modello mappa la maturità dell'organizzazione, restituendo tramite radar chart una diagnosi chiara delle vulnerabilità e delle loro interdipendenze.



Riferimenti bibliografici

1. Beck, U. (1992). *Risk Society: Towards a New Modernity*. Sage Publications. (Opera originale pubblicata nel 1986).
2. Benjamin, R. (2019). *Race After Technology: Abolitionist Tools for the New Jim Code*. Polity.
3. Bolukbasi, T., Chang, K. W., Zou, J. Y., Saligrama, V., & Kalai, A. T. (2016). Man is to computer programmer as woman is to homemaker? Debiasing word embeddings. *Advances in Neural Information Processing Systems*, 29, 4349-4357.
4. Broussard, M. (2018). *Artificial Unintelligence: How Computers Misunderstand the World*. MIT Press.
5. Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of the 1st Conference on Fairness, Accountability and Transparency (FAcT)*, 77-91.
6. Campolo, A., & Crawford, K. (2020). Enchanted Determinism: Power without Responsibility in Artificial Intelligence. *Engaging Science, Technology, and Society*, 6, 1-19.
7. Cheatham, B., Javanmardian, K., & Samandari, H. (2019). Confronting the risks of artificial intelligence. *McKinsey Quarterly*.
8. Costanza-Chock, S. (2020). *Design Justice: Community-Led Practices to Build the Worlds We Need*. MIT Press.
9. Crawford, K. (2021). *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. Yale University Press.
10. Dastin, J. (2018). Amazon scraps secret AI recruiting tool that showed bias against women. *Reuters*, 10(9).
11. De Gregorio, G. (2023). Il potere normativo dell'intelligenza artificiale. *CGSL Working Papers*, 4/2023, Catolica Global School of Law.
12. Di Cataldo L. (2025). La transizione intelligente: la produzione industriale e il settore sanitario nel XXI secolo. *Neural Nexus Review*, 1: 34-77.
13. D'Ignazio, C., & Klein, L. F. (2020). *Data Feminism*. MIT Press.
14. Eubanks, V. (2018). *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. St. Martin's Press.
15. Floridi, L., Cows, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., ... & Vayena, E. (2018). AI4People—An ethical framework for a good AI society. *Minds and Machines*, 28(4), 689-707.
16. Gebru, T. (2020). Race and Gender. In M. D. Dubber, F. Pasquale, & S. Das (Eds.), *The Oxford*



Handbook of Ethics of AI. Oxford University Press.

17. Gibin, M., Krasanakis, E., & Grumodas, K. (2025). Negoziare l'equità dell'IA: un appello a riequilibrare i rapporti di potere. *AI & Society*, 1-18.
18. Kitchin, R. (2017). Thinking critically about and researching algorithms. *Information, Communication & Society*, 20(1), 14-29.
19. Kuźmicz, M. M. (2025). Balancing the scales: equilibrium and power relations in the AI era. *AI & Society*, 40, 5195-5212.
20. Maas, J. (2022). Machine learning and power relations. *AI & Society*, 38, 1493-1500.
21. Michaels, A. C. (2020). Artificial Intelligence, Legal Change, and Separation of Powers. *University of Cincinnati Law Review*, 88(4), 1083-1104.
22. Müller, V. C. (2016). Editorial: Risks of general artificial intelligence. *Journal of Experimental & Theoretical Artificial Intelligence*, 26(3), 297-301.
23. Noble, S. U. (2018). *Algorithms of Oppression: How Search Engines Reinforce Racism*. NYU Press.
24. Nuzzo, A. (2019). Algoritmi e regole. *Analisi Giuridica dell'Economia*, 1, 1-15.
25. O'Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown.
26. Pasquale, F. (2015). *The Black Box Society: The Secret Algorithms That Control Money and Information*. Harvard University Press.
27. Schuett, J. (2024). Risk management in the Artificial Intelligence Act. *European Journal of Risk Regulation*, 15(2), 367-385.
28. Wajcman, J. (2004). *TechnoFeminism*. Polity Press.
29. West, S. M., Whittaker, M., & Crawford, K. (2019). *Discriminating Systems: Gender, Race and Power in AI*. AI Now Institute.
30. Yudkowsky, E. (2008). Artificial Intelligence as a Positive and Negative Factor in Global Risk. In N. Bostrom & M. M. Ćirković (Eds.), *Global Catastrophic Risks* (pp. 308-345). Oxford University Press.