

DATI E ALGORITMI TRA DISUGUAGLIANZE E RESISTENZE

a cura di
Stefania Tusini, Davide Bennato,
Maria Dentale

FrancoAngeli 

INFERENZE

EVIDENZE

INFERENZE

EVIDENZE

Inferenze/Evidenze

collana diretta da *Antonio Fasanella e Carmelo Lombardo*

Comitato scientifico: Maria Carmela Agodi (Università degli studi di Napoli Federico II), Giuseppe Anzera (Sapienza Università di Roma), Adele Bianco (Università degli studi "G. d'Annunzio" Chieti-Pescara), Christian Borch (Copenhagen Business School), Andrea Borghini (Università di Pisa), Wayne Brekhus (University of Missouri/Columbia), Ernesto D'Albergo (Sapienza Università di Roma), Alessandra Decataldo (Università degli studi di Milano Bicocca), Maria Paola Faggiano (Sapienza Università di Roma), Giovanna Gianturco (Sapienza Università di Roma), Srebranka Letina (University of Glasgow), Mariano Longo (Università del Salento), Veronica Lo Presti (Sapienza Università di Roma), Krzysztof T. Konecki (University of Łódź), Alberto Marinelli (Sapienza Università di Roma), Stefano Nobile (Sapienza Università di Roma), Paolo Parra Saiani (Università di Genova), Massimo Pendenza (Università degli studi di Salerno), Olli Pyyhtinen (University of Tampere), Lorenzo Sabetta (Sapienza Università di Roma), Hizky Shoham (Bar-Ilan University), Stefania Tusini (Università per Stranieri di Perugia), Dieter Vandebroeck (Free University of Brussels), Petri Ylikoski (University of Helsinki).

Comitato editoriale: Lorenzo Barbanera (Sapienza Università di Roma), Ernesto Dario Calò (Sapienza Università di Roma), Michela Cavagnuolo (Università degli Studi di Roma "Foro Italico"), Maria Dentale (Università per Stranieri di Perugia), Raffaella Gallo (Sapienza Università di Roma), Melissa Mongiardo (Università della Toscana).

Inferenze/Evidenze intende promuovere il pluralismo delle idee e un approccio integrato di teoria e ricerca, configurandosi come uno spazio di condivisione di prospettive concettuali, strategie di indagine ed esperienze empiriche centrate su un'ampia varietà di temi e problemi tipici del mondo contemporaneo. Guarda a percorsi investigativi capaci di valorizzare la pratica dell'immaginazione sociologica e, attraverso disegni di ricerca rigorosi e innovativi, ancorati a strutture teoriche e a sufficienti e controllate basi di dati, di favorire il più possibile lo sviluppo di programmi di ricerca pluralistici e integrati.

I volumi pubblicati sono sottoposti alla valutazione anonima di almeno due *referees* esperti.

FrancoAngeli
OPEN  ACCESS



Il presente volume è pubblicato in open access, ossia il file dell'intero lavoro è liberamente scaricabile dalla piattaforma **FrancoAngeli Open Access** (<http://bit.ly/francoangeli-oa>).

FrancoAngeli Open Access è la piattaforma per pubblicare articoli e monografie, rispettando gli standard etici e qualitativi e la messa a disposizione dei contenuti ad accesso aperto. Oltre a garantire il deposito nei maggiori archivi e repository internazionali OA, la sua integrazione con tutto il ricco catalogo di riviste e collane FrancoAngeli massimizza la visibilità, favorisce facilità di ricerca per l'utente e possibilità di impatto per l'autore.

Per saperne di più: [Pubblica con noi](#)

I lettori che desiderano informarsi sui libri e le riviste da noi pubblicati possono consultare il nostro sito Internet: www.francoangeli.it e iscriversi nella home page al servizio "[Informatemi](#)" per ricevere via e-mail le segnalazioni delle novità.

DATI E ALGORITMI TRA DISUGUAGLIANZE E RESISTENZE

a cura di
Stefania Tusini, Davide Bennato,
Maria Dentale

FrancoAngeli 

INFERENZE

EVIDENZE

Il volume è stato finanziato nell'ambito del progetto di ricerca "GILDA – Gender data bias, femminismo e disuguaglianze nell'era digitale", CUP D95F21002910001, finanziato dall'Unione Europea – NextGenerationEU, in attuazione del D.M. MUR n. 737/2021 "Criteri di riparto e utilizzazione del Fondo per la promozione e lo sviluppo delle politiche del Programma Nazionale per la Ricerca (PNR) 2021–2027".

Isbn: 9788835185925

Isbn e-book Open Access: 9788835195153

Copyright © 2026 by FrancoAngeli s.r.l., Milano, Italy.

Publicato con licenza *Creative Commons*
Attribuzione-Non Commerciale-Non opere derivate 4.0 Internazionale
(CC-BY-NC-ND 4.0).

Sono riservati i diritti per Text and Data Mining (TDM), AI training e tutte le tecnologie simili.

L'opera, comprese tutte le sue parti, è tutelata dalla legge sul diritto d'autore.
L'Utente nel momento in cui effettua il download dell'opera accetta tutte le condizioni della licenza d'uso dell'opera previste e comunicate sul sito
<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.it>

Gli eventuali link attivi e QR code inseriti nel volume sono forniti dall'autore. L'editore non si assume alcuna responsabilità sui link attivi e QR code ivi contenuti che rimandano a siti non appartenenti a FrancoAngeli.

L'autore dichiara che, nella fase di stesura del presente volume,
non ha utilizzato strumenti di IA generativa.

Indice

Introduzione , di Stefania Tusini, Davide Bennato, Maria Dentale	pag.	9
1. Dati e algoritmi: parole chiave della contemporaneità	»	9
2. Articolazione del volume	»	15
Riferimenti bibliografici	»	18
1. La società algomorfica: dispositivi convergenti, agentività situata e co-produzione del sapere , di Edmondo Grassi	»	21
1. Frattura epistemica e legittimazione del sapere	»	21
2. Soglie e dispositivi: genealogia minima della società algomorfica	»	26
3. Agentività situata: una griglia analitica sui casi studio	»	31
4. Oltre il binomio morale: responsabilità distribuita e politica del possibile	»	39
Riferimenti bibliografici	»	41
2. Pregiudizi computazionali. Il fenomeno dei bias tra algoritmi e politica , di Davide Bennato	»	44
1. Introduzione	»	44
2. Algoritmi e pregiudizi: i termini del dibattito	»	45
3. Intelligenza artificiale e pregiudizi	»	48
4. Il bias come problema?	»	51
5. Come gestire i bias fra soluzione e persistenza	»	52
6. Conclusioni: affrontare bias e discriminazioni evitando il mito della neutralità	»	55
Riferimenti bibliografici	»	56

3. Algoritmi e Big Data: Un’analisi critica dei meccanismi di riproduzione delle disuguaglianze sociali , di Fiorenza Deriu	pag.	58
1. Introduzione	»	58
2. Ambiti in cui si manifestano le diverse forme di disuguaglianza	»	60
3. La riproduzione delle disuguaglianze sociali. Alcuni studi di caso ed esperimenti	»	65
4. Osservazioni conclusive: verso un approccio sociologico alla comprensione del fenomeno	»	75
Riferimenti bibliografici	»	77
4. Discriminazione algoritmica e disuguaglianze sociali: le nuove sfide della Valutazione , di Maria Dentale	»	79
1. La relazione tra <i>data science</i> e Valutazione. Un’occasione mancata?	»	79
2. Sistemi <i>data-intensive</i> e Valutazione. Note critiche su un rapporto <i>in fieri</i>	»	81
3. Valutazione dei sistemi <i>data-intensive</i> e teoria della complessità	»	85
4. <i>Bias, black box</i> e meccanismi sociali: esempi di discriminazione algoritmica	»	89
5. Per concludere: <i>matters of fact</i> or <i>matters of concern?</i>	»	95
Riferimenti bibliografici	»	96
5. Il <i>data journalism</i> tra pretesa di “verità” e controversia politico-culturale , di Rolando Marini, Matteo Gerli	»	100
1. Il rapporto tra scienza, dati e giornalismo: una ricognizione	»	100
2. Genealogia del <i>data journalism</i>	»	104
3. Il <i>data journalism</i> tra conoscenza aperta e controversia interpretativa	»	108
4. La costruzione <i>data-driven</i> del tema immigrazione nel giornalismo italiano	»	112
Riferimenti bibliografici	»	119

6. Pensare, sentire, agire l’algoritmo. Un’analisi qualitativa della consapevolezza algoritmica, di Enrica Amaturò, Ciro Clemente De Falco, Caterina Ambrosio	pag.	124
1. Introduzione	»	124
2. Studiare la consapevolezza algoritmica	»	126
2.1. Modelli concettuali della consapevolezza algoritmica	»	126
2.2. Approcci empirici e modalità di restituzione dei risultati	»	128
3. Disegno della ricerca e strategia di campionamento	»	131
4. Dimensioni analitiche e costruzione della tipologia	»	132
4.1. Articolazione empirica delle dimensioni analitiche	»	132
4.2. Dalle categorie ai tipi analitici	»	135
5. Riflessioni conclusive	»	137
Riferimenti bibliografici	»	139
7. “<i>Palestinian cooking is resistance</i>”. <i>Algorithmic activism e foodie culture nella resistenza digitale palestinese</i>, di Claudia Cantale	»	142
1. Introduzione	»	142
2. Attivismo digitale: breve revisione della letteratura	»	143
3. Attivismo e resistenza digitale nella causa palestinese	»	145
4. La cucina palestinese è <i>resistenza</i>	»	148
5. Il disegno della ricerca	»	150
6. Riflessioni conclusive	»	156
Riferimenti bibliografici	»	163
Le autrici e gli autori	»	167

Introduzione

di *Stefania Tusini, Davide Bennato, Maria Dentale**

1. Dati e algoritmi: parole chiave della contemporaneità

Il titolo che abbiamo scelto per questo volume tiene insieme quattro termini – *dati, algoritmi, disuguaglianze, resistenze* – che a prima vista potrebbero sembrare appartenenti a universi distinti. E invece, è proprio nel loro incontro che si gioca una delle partite più delicate e importanti del nostro tempo.

Parleremo di *dati*: il nuovo “petrolio”, secondo una metafora estrattiva che svela la connotazione ideologica del processo, ma anche – e soprattutto – “la nuova grammatica” attraverso cui si costruisce la realtà sociale, politica, economica. Dati, dunque, come nuova infrastruttura del mondo: raccolti in tempo reale, organizzati in flussi, utilizzati per analizzare comportamenti, prevedere azioni, indirizzare decisioni pubbliche e private. Senza dimenticare che i dati sono il prodotto di scelte – su cosa rilevare, come raccogliere, cosa ignorare – e queste scelte non sono mai neutre perché riflettono visioni del mondo, priorità, poteri.

Il volume tratterà poi di *algoritmi*, impiegati ogni giorno per orientarci nel traffico, per cercare informazioni, per decidere cosa leggere, acquistare o guardare. Ma che vengono anche utilizzati per classificare, prevedere, selezionare: chi ha diritto a un servizio, chi è considerato a rischio, chi è meritevole di attenzione. Algoritmi come dispositivi che decidono o influenzano una quantità crescente di aspetti della nostra vita: dalla pubblicità personalizzata alle diagnosi mediche, dalla selezione del personale, fino alla distribuzione delle risorse in ambito educativo, sanitario o sociale. Gli algoritmi promettono imparzialità, efficienza, oggettività, ma sempre più spesso scopriamo che possono riprodurre – e talvolta amplificare – pregiudizi, disuguaglianze, discriminazioni.

* Il capitolo è il risultato della riflessione congiunta delle autrici e dell'autore.

Si parlerà degli effetti che questi strumenti possono avere, degli squilibri che producono o amplificano, delle asimmetrie di potere che rendono visibili alcuni soggetti e invisibili altri, delle resistenze possibili, individuali e collettive, che possono emergere come forma di risposta critica, etica, politica. Infatti, come detto, accanto a questi pilastri della modernità digitale (dati e algoritmi), abbiamo scelto di accostare due termini che ci riportano alla dimensione del conflitto e della possibilità: *disuguaglianze e resistenze*.

Le prime (le disuguaglianze) sono ormai ben documentate da decine di studi (Zajko, 2022; Alegria e Yeh, 2023; Deriu, 2026) che mostrano come le tecnologie digitali non solo non eliminino le disuguaglianze sociali, ma talvolta le rendano più sottili, più profonde, più difficili da smascherare. Le altre, le resistenze, sono le pratiche individuali e collettive con cui soggetti diversi – cittadini, accademici, attivisti, tecnologi, artisti – cercano di rispondere a questi squilibri, immaginando e costruendo alternative.

Non si tratta, ovviamente, di demonizzare i dati o gli algoritmi, né tantomeno di evocare scenari distopici, soprattutto per il rischio (concreto) di vedersi superati dalla realtà. Il nostro obiettivo è piuttosto porre domande scomode, che troppo spesso rimangono ai margini del dibattito pubblico e della formazione accademica. Come nascono le disuguaglianze nei dati? Chi decide quali dati contano, e quali no? Quali soggettività sono rappresentate, e quali restano invisibili? Come funzionano gli algoritmi che determinano l'accesso a un lavoro, a un'assistenza sanitaria, a un contenuto informativo? In che modo queste tecnologie possono rafforzare stereotipi di genere, razza, classe, abilità? E ancora: di fronte a tutto questo, quali margini di manovra abbiamo? Come possiamo costruire delle resistenze che non siano solo difensive, ma anche generative? Come possiamo immaginare una “giustizia algoritmica”, una governance democratica dei dati, un'etica che non sia solo posticcia, ma intrinseca ai processi di progettazione?

In un'epoca in cui si parla molto di intelligenza artificiale, spesso con toni entusiastici o, al contrario, apocalittici, questo volume intende proporre una riflessione lucida e complessa. Non interessa né aderire alla retorica tecno-soluzionista né cedere a una posizione tecnofobica. Interessa invece capire come vengono prodotti i dati, quali logiche guidano la costruzione degli algoritmi, chi trae beneficio e chi invece rischia di essere escluso o penalizzato da questi sistemi.

Facciamo alcuni esempi dato che, come è noto, ogni strumento (e gli algoritmi lo sono) può essere usato, al netto delle affordances, costruttivamente ma anche distruttivamente. Come si intuisce, i settori di applicazione sono infiniti. Si pensi all'uso dell'intelligenza artificiale nel campo della ricerca biomedica grazie alla quale è possibile processare enormi quantità di dati al fine, ad esempio, di individuare antibiotici efficaci verso i batteri

antibiotico-resistenti, come hanno fatto alcuni scienziati del MIT (Wong *et al.*, 2024). Oppure impiegare l'intelligenza artificiale per rendere più rapidi e precisi i modelli meteorologici a fini previsionali (Kotsuki *et al.*, 2025). Pensiamo, ancora, all'area delle scoperte basate sulla metodica della *systematic review* che consiste nel setacciare sistematicamente tutta la letteratura scientifica esistente su un dato tema con la finalità di trovare collegamenti, idee, nuove ipotesi e/o immaginare eventuali nuovi esperimenti.

La frontiera, in questo campo, sembra essere quella dei cosiddetti “laboratori autonomi” in cui scienziati robot useranno l'intelligenza artificiale per testare nuove ipotesi essendo in grado di realizzare migliaia di esperimenti e di sviluppare teorie inaspettate senza l'intervento umano.

Un altro campo particolarmente interessante è quello del cosiddetto *peacetech*, ovvero l'impiego dell'intelligenza artificiale e delle tecnologie digitali al servizio della diplomazia, della prevenzione dei conflitti e della costruzione della pace. In questo ambito, l'IA viene utilizzata per analizzare enormi quantità di dati eterogenei – dai flussi economici e migratori ai contenuti dei social media, dalle immagini satellitari alle informazioni climatiche – al fine di individuare segnali precoci di instabilità politica o di escalation violenta. Attraverso modelli predittivi e sistemi di *early warning*, queste tecnologie possono aiutare organizzazioni internazionali, governi e ONG a intervenire prima che un conflitto esploda, orientando strategie diplomatiche, missioni di mediazione o politiche di prevenzione mirate.

Il *peacetech*, tuttavia, non si limita alla previsione dei conflitti. Include anche strumenti per facilitare il dialogo tra attori ostili, supportare processi di negoziazione complessi, contrastare la disinformazione e l'incitamento all'odio online, che spesso alimentano tensioni sociali e politiche. In questo senso, l'IA può diventare un'infrastruttura di supporto alla diplomazia, capace di ampliare la comprensione dei contesti locali e di rendere più informate e tempestive le decisioni politiche.

Al tempo stesso, il ricorso all'intelligenza artificiale in un ambito così delicato solleva questioni etiche e politiche rilevanti: chi definisce i parametri del rischio? Quali dati vengono considerati affidabili? E come evitare che strumenti pensati per la pace vengano riutilizzati per il controllo, la sorveglianza o la repressione? Proprio per questo, il *peacetech* rappresenta un terreno cruciale in cui si intrecciano tecnologia, potere e responsabilità, e in cui l'uso dell'IA richiede non solo competenze tecniche, ma anche una forte cornice normativa, trasparente e orientata ai diritti umani.

Esistono tecnologie, come detto, in grado di prevedere potenziali aree di escalation; o di costruire modelli complessi del comportamento di intere popolazioni. Sono strumenti in grado di realizzare *opinion mapping* individuando aree di consenso all'interno di comunità divise. Impiegare sistemi

di *opinion mapping* significa infatti utilizzare strumenti di analisi computazionale per mappare le posizioni, le narrazioni e le sensibilità presenti all'interno di una comunità, soprattutto quando essa è attraversata da forti divisioni politiche, culturali o identitarie. Attraverso l'analisi di grandi quantità di dati testuali – come discussioni sui social media, forum, documenti pubblici o interventi nei media – questi sistemi permettono di individuare non solo le linee di conflitto più evidenti, ma anche aree di sovrapposizione e consenso latente che spesso rimangono invisibili nel dibattito pubblico polarizzato.

In questo senso, l'*opinion mapping* non serve a semplificare o appiattire le differenze, ma a renderle leggibili. Visualizzare lo “spazio delle opinioni”, rende possibile comprendere come le posizioni si distribuiscano, quali temi fungano da catalizzatori del conflitto e quali, al contrario, rappresentino potenziali terreni comuni. Questo può risultare particolarmente utile in contesti di mediazione, negoziazione o progettazione di politiche pubbliche, dove identificare punti di accordo, anche parziali o fragili, può aprire spazi di dialogo altrimenti bloccati.

Tuttavia, l'uso dell'*opinion mapping* richiede cautela. Le opinioni non sono entità statiche né riducibili a semplici cluster: sono situate, mutevoli e spesso ambivalenti. Per questo, tali strumenti dovrebbero essere intesi come supporti interpretativi, non come dispositivi decisionali automatici. Se impiegati in modo responsabile e trasparente, possono contribuire a spostare l'attenzione dalla mera contrapposizione alla ricerca di convergenze, offrendo alla diplomazia, alla governance e alla costruzione della pace una base informativa più ricca e articolata.

Accanto ad impieghi virtuosi come quelli su descritti, dobbiamo però documentare anche un uso non propriamente pacifico, come nel caso di *Lavender* (+972 Magazine, 2024), un sistema di IA impiegato dai militari israeliani dell'IDF nel conflitto contro i palestinesi a cui è stato dato il compito di individuare target bersaglio sulla base di alcuni criteri quali parentela con membri di organizzazioni militanti (Hamas o Jihad), residenza in determinate aree, contatti telefonici con sospetti militanti, e così via.

Il suo scopo è analizzare grandi quantità di dati – come comunicazioni, movimenti, profili e reti sociali – per generare liste di individui considerati potenziali membri di organizzazioni armate. In questo modo *Lavender* a quanto pare avrebbe individuato circa 37.000 obiettivi e indicato la localizzazione migliore per ucciderli (*The Guardian*, 2024). Secondo inchieste giornalistiche (*Ibidem*; Valori, 2024), è la stessa IA che stabilisce il margine di tolleranza rispetto alle perdite civili accettabili e sembra sia considerata ammissibile la morte di 15-20 civili per un bombardamento destinato a uccidere un miliziano di basso rango, e fino a 100 civili per uno di alto ran-

go. Inoltre, ancora l'intelligenza artificiale applicata ad un altro sistema denominato *Gospel* indica i luoghi in cui colpire i bersagli, scegliendo quelli che garantiscono più di altri l'eliminazione del nemico. *Gospel* normalmente suggerisce di bombardare i bersagli non in strutture più difficili da attaccare come i tunnel e altri luoghi protetti, ma a casa loro, senza alcuno scrupolo per le vittime collaterali, usando *dumb bomb* che colpiscono indiscriminatamente tutti gli abitanti dell'abitazione individuata.

L'evoluzione tecnologica ha trasformato i campi di battaglia in laboratori a cielo aperto, dove il codice conta quanto il piombo. Come sopra illustrato, diverse sono le tecnologie che contribuiscono agli scenari di guerra contemporanei, anche se quella maggiormente sotto osservazione, sia per l'impatto operativo che per delicate implicazioni politiche, è senza dubbio *Palantir Technologies*.

Fondata da Peter Thiel e Alex Karp, Palantir non produce armi fisiche, ma "motori" algoritmici capaci di processare volumi sterminati di dati eterogenei per prevedere crimini, rintracciare individui e ottimizzare i target militari. La sua ascesa è indissolubilmente legata al sodalizio con l'amministrazione Trump, periodo in cui l'azienda ha consolidato il proprio ruolo di pilastro tecnologico della sicurezza nazionale. Thiel, uno dei pochi sostenitori della Silicon Valley a supportare apertamente Donald Trump nel 2016, ha agito come ponte tra il governo e l'industria tech. Sotto quella presidenza, Palantir ha ottenuto contratti cruciali, non ultimo quello con l'ICE (Immigration and Customs Enforcement) per la gestione dei flussi migratori e delle espulsioni. Tale vicinanza ha sollevato interrogativi etici profondi: quando il potere politico si fonde con una capacità di calcolo quasi onnisciente, il confine tra sicurezza collettiva e controllo autoritario diventa pericolosamente sottile, trasformando l'algoritmo in un arbitro invisibile di vita e di morte.

Tutto questo aiuta anche a capire il motivo per cui questa guerra abbia un bilancio di vittime civili così elevato e un livello di distruzione delle abitazioni e delle infrastrutture senza precedenti.

In questo contesto, crescentemente complesso, crediamo che l'Università abbia un ruolo decisivo. Non è possibile limitarsi a formare esperti di dati o di informatica: è necessario invece istruire cittadini critici, ricercatori consapevoli, professionisti capaci di tenere insieme rigore tecnico e sensibilità sociale.

Per farlo al meglio, sarebbe auspicabile superare le barriere disciplinari e costruire un dialogo tra sociologia e data science, tra filosofia e ingegneria, tra diritto e antropologia, tra medicina e studi di genere. Solo attraverso un sapere ibrido, contaminato, aperto, possiamo sperare di comprendere – e forse anche orientare – i processi in atto.

Questo volume prova a dare il suo contributo, raccogliendo analisi teoriche, presentazioni di ricerche empiriche, discussioni di casi concreti, proposte operative. Ci auguriamo che da questo confronto possano nascere non solo nuove conoscenze, ma anche nuovi percorsi di formazione e nuovi strumenti di intervento.

Oggi più che mai è importante riaffermare un principio che sembra a volte dimenticato: la tecnica non è neutra, e ogni scelta tecnica è anche una scelta politica perché contribuisce a definire come il potere viene esercitato e distribuito all'interno di una società. La selezione di un parametro, la definizione di un limite, la costruzione di una categoria – tutti questi atti hanno delle conseguenze.

Affermare che la tecnica non è neutrale significa infatti riconoscere che ogni artefatto tecnologico incorpora visioni del mondo, valori e rapporti di forza. Le tecnologie non si limitano a eseguire funzioni: strutturano possibilità d'azione, orientano comportamenti, rendono alcune scelte più facili e altre più difficili.

Questo vale in modo particolarmente evidente per i sistemi algoritmici e di intelligenza artificiale. La selezione di un parametro, la definizione di una soglia di rilevanza o di rischio, la costruzione di una categoria analitica non sono operazioni puramente tecniche: sono atti interpretativi che traducono realtà complesse in modelli semplificati. Stabilire, ad esempio, cosa sia un “comportamento anomalo”, “contenuto accettabile” o “profilo ad alto rischio” implica sempre una presa di posizione su ciò che è considerato normale, desiderabile o legittimo.

Queste scelte hanno conseguenze materiali e simboliche. Determinano chi viene incluso e chi escluso da un servizio, da un beneficio o da uno spazio di visibilità; chi viene premiato e chi penalizzato; chi ottiene riconoscimento e chi, al contrario, resta marginalizzato o silenziato. Spesso tali effetti colpiscono in modo sproporzionato gruppi già vulnerabili, amplificando disuguaglianze preesistenti sotto una veste apparentemente oggettiva e automatica.

Riconoscere la dimensione politica della tecnica non significa rifiutarla, ma assumerne la responsabilità. Significa rendere esplicite le scelte incorporate nei sistemi tecnologici, aprirle al dibattito pubblico, e interrogarsi su quali valori esse promuovano e su quali interessi servano. Solo in questo modo la progettazione tecnica può diventare uno spazio di deliberazione democratica, anziché un ambito opaco sottratto al confronto collettivo.

Per questo è fondamentale insistere su tre parole chiave: *potere*, perché ogni sistema algoritmico è anche un dispositivo di potere, che distribuisce risorse, visibilità, possibilità. Ogni sistema algoritmico, si è detto, non è mai neutrale: è un dispositivo che incorpora decisioni politiche, economi-

che e culturali, e che agisce concretamente sulla distribuzione di risorse, attenzione, opportunità e diritti. Gli algoritmi stabiliscono chi viene visto e chi resta invisibile, quali comportamenti vengono premiati e quali penalizzati. In questo senso, governare gli algoritmi significa interrogare le forme di potere che essi esercitano e le asimmetrie che producono o rafforzano.

La *trasparenza* è altrettanto cruciale, perché senza accesso ai codici, ai dati, alle logiche sottostanti, ogni forma di controllo sociale diventa impossibile; senza la possibilità di comprendere come funzionano questi sistemi – quali dati utilizzano, secondo quali criteri prendono decisioni, quali interessi li orientano – ogni tentativo di controllo democratico risulta svuotato. L’opacità algoritmica non è solo un problema tecnico, ma un problema politico: rende le decisioni automatiche incontestabili, naturali, inevitabili. Rendere trasparenti codici, dataset e logiche operative non significa soltanto “aprire” le scatole nere, ma creare le condizioni affinché cittadini, ricercatori e istituzioni possano discutere, contestare e trasformare questi sistemi.

Infine, l’*immaginazione*, perché non si può limitarsi a criticare ciò che esiste, ma anche essere capaci di progettare ciò che potrebbe essere: altri modi di raccogliere i dati, di trattarli, di condividerli, di governarli. Criticare l’esistente è necessario, ma insufficiente se non è accompagnato dalla capacità di pensare alternative. Immaginare significa rifiutare l’idea che gli attuali modelli di raccolta, trattamento e sfruttamento dei dati siano gli unici possibili. Significa progettare altri ecosistemi digitali: forme di governance collettiva dei dati, infrastrutture pubbliche o cooperative, algoritmi orientati al bene comune piuttosto che all’estrazione di valore. Senza immaginazione politica e tecnologica, la critica rischia di restare sterile; con essa, diventa invece un atto generativo, capace di aprire spazi per futuri diversi.

In conclusione, e di conseguenza, il nostro augurio è che questo volume sia non solo un’occasione di apprendimento, ma anche di messa in discussione, di condivisione, di connessione tra saperi e pratiche. Siamo convinti che ci sia bisogno di strumenti come questo, in cui la riflessione accademica si fa spazio pubblico, in cui il sapere si apre alla responsabilità e alla trasformazione, e possa ispirare ricerche future, percorsi formativi nuovi, e – perché no – anche politiche più attente, più giuste, più inclusive.

2. Articolazione del volume

Il volume si articola in sette capitoli, i quali affrontano il rapporto tra *dati*, *algoritmi*, *disuguaglianze* e *resistenze* da angolature differenti e ruotano intorno ai tre nodi analitici richiamati in precedenza: le forme del *potere algoritmico*; le condizioni – e i limiti – della *trasparenza*; le possibilità di

resistenza e di *immaginazione* alternativa. Mettendo in luce tensioni e ambivalenze che attraversano le pratiche algoritmiche e mostrando come gli stessi strumenti di IA possano assumere, a seconda dei contesti e degli usi, funzioni divergenti (efficienza e razionalizzazione vs esclusione, opacizzazione e riproduzione delle disuguaglianze), ciascun saggio contribuisce a rendere visibile la linea di frattura che separa la “promessa di neutralità algoritmica” dagli “effetti sociali” concreti delle decisioni automatizzate, spesso occultata da narrazioni semplificanti o tecnicistiche.

Il volume si apre con il contributo di Edmondo Grassi, “La società algomorfica: dispositivi convergenti, agentività situata e co-produzione del sapere”, che offre una cornice teorica per leggere il potere algoritmico come forma emergente di governo delle pratiche, delle soggettività e dei regimi di conoscenza. Attraverso la nozione di “soglia algomorfica”, il contributo mostra come l’integrazione della GenAI non rappresenti soltanto un avanzamento tecnico, ma una frattura epistemica che riconfigura le condizioni di legittimazione del sapere e le forme dell’agentività. I casi discussi – dai robot umanoidi alle piattaforme ludiche, dai *lovebot* alle quasi-soggettività digitali – mostrano come potere computazionale e soggettività si co-producano attraverso routine, vincoli infrastrutturali e attribuzioni di agentività.

Il tema del potere algoritmico, che si traduce in processi di classificazione, esclusione e discriminazione sociale, è affrontato in modo più diretto nel contributo di Davide Bennato, “Pregiudizi computazionali. Il fenomeno dei bias tra algoritmi e politica”. Qui l’attenzione si concentra sui bias algoritmici come esito di processi sociali, culturali e decisionali che precedono la computazione. Il capitolo ricostruisce le principali fonti dei pregiudizi incorporati nei sistemi algoritmici e ne discute gli effetti in ambiti ad alto impatto sociale, mostrando come il dataismo e l’illusione di neutralità trasformino scelte politiche e morali in output apparentemente oggettivi. La distinzione proposta tra “deviazioni statistiche inevitabili” e “discriminazioni effettive” consente di spostare il fuoco dalla sola “correzione tecnica” alla responsabilità sociale e politica dei sistemi decisionali automatizzati.

Una riflessione affine, ma più esplicitamente centrata sui meccanismi di riproduzione delle disuguaglianze sociali, è sviluppata nel contributo di Fiorenza Deriu, “Algoritmi e Big Data: un’analisi critica dei meccanismi di riproduzione delle disuguaglianze sociali”. Il capitolo mostra come l’intelligenza artificiale agisca da potente amplificatore di asimmetrie preesistenti, intrecciandosi con narrazioni dominanti, assetti oligopolistici e infrastrutture opache. Attraverso i concetti di capitalismo della sorveglianza, governamentalità algoritmica e *black box society*, il contributo chiarisce come la normalizzazione delle disuguaglianze passi anche dalla depoliticizzazio-

ne della tecnologia e dall'occultamento delle relazioni di potere che ne informano progettazione e utilizzo. In questo quadro, la resistenza è intesa come lavoro critico di decostruzione delle narrazioni e come progetto interdisciplinare orientato alla giustizia sociale e a forme sostanziali – non solo formali – di trasparenza.

Il confronto con l'opacità algoritmica nei processi decisionali pubblici è al centro del contributo di Maria Dentale, "Discriminazione algoritmica e diseguaglianze sociali: le nuove sfide della Valutazione". Il capitolo analizza il ruolo crescente di big data e intelligenza artificiale nei processi valutativi, mostrando come l'automazione possa comprimere i margini di discrezionalità e tradurre disuguaglianze storiche in criteri operativi standardizzati. Attraverso casi emblematici di allocazione algoritmica delle risorse pubbliche, il contributo mostra come la Valutazione possa assumere un ruolo cruciale nel rendere intelligibili e discutibili i meccanismi sociali incorporati nei modelli *data-intensive*. In questo senso, la resistenza non coincide con il rifiuto degli strumenti computazionali, ma con l'adozione di una postura riflessiva e metodologicamente ibrida, capace di integrare analisi algoritmica e teoria dell'azione, con un'attenzione ai contesti organizzativi, istituzionali e relazionali in cui tali strumenti vengono progettati e utilizzati.

Il rapporto tra dati, trasparenza e costruzione pubblica della realtà è esplorato nel contributo di Rolando Marini e Matteo Gerli, "Il data journalism tra pretesa di "verità" e controversia politico-culturale", dedicato al nesso tra scienza, dati e giornalismo. Attraverso una ricostruzione genealogica e l'analisi di casi tratti dal giornalismo contemporaneo, il capitolo mostra come l'uso dei dati non neutralizzi il conflitto interpretativo, ma diventi esso stesso terreno di contesa. I dati, lungi dall'essere evidenze innocue, operano come risorse discorsive mobilitate all'interno di rapporti di potere che strutturano visibilità, legittimità e accesso allo spazio pubblico. La trasparenza, in questo quadro, non coincide con la semplice disponibilità dei dati, ma con la possibilità di interpretarli criticamente e di inserirli entro cornici di senso socialmente e politicamente rilevanti.

Le forme di resistenza algoritmica situata sono al centro del contributo di Enrica Amatore, Ciro Clemente De Falco e Caterina Ambrosio, "Pensare, sentire, agire l'algoritmo", che propone una lettura della consapevolezza algoritmica come dimensione relazionale, situata e processuale. A partire da una ricerca qualitativa che ha coinvolto diversi attivisti del web, il capitolo mette in discussione approcci che riducono la "consapevolezza algoritmica" a competenza misurabile, mostrando come adattamento, critica, elusione e rifiuto degli algoritmi configurino modalità differenti – spesso ambivalenti – di negoziazione tra visibilità, agency e condizioni materiali dell'azione collettiva. In questa prospettiva, la resistenza emerge come in-

sieme di pratiche situate che si sviluppano all'interno dei vincoli imposti dalle piattaforme.

Il volume si chiude con il contributo di Claudia Cantale, "Palestinian cooking is resistance. Algorithmic activism e foodie culture nella resistenza digitale palestinese", che porta l'attenzione su pratiche vernacolari di immaginazione politica. Attraverso l'analisi dell'attivismo digitale palestinese su Instagram, il capitolo mostra come contenuti legati al cibo e alla quotidianità diventino dispositivi di contro-narrazione, memoria e decolonizzazione simbolica. Le strategie di *algorithmic amplification, evasion e hijacking* analizzate nel contributo evidenziano come la resistenza possa operare dall'interno delle piattaforme, trasformando linguaggi apparentemente marginali in strumenti politici capaci di produrre visibilità, empatia e riconoscimento.

Nel suo insieme, il volume restituisce la complessità dei processi attraverso cui dati e algoritmi entrano a far parte della vita sociale, contribuendo a ridisegnare confini, gerarchie e possibilità di azione. Le disuguaglianze che ne derivano non sono effetti inevitabili, ma il risultato di configurazioni sociotecniche specifiche, che possono essere analizzate, discusse e, in alcuni casi, trasformate. È in questo spazio di "indeterminazione" che si collocano le pratiche di resistenza e di immaginazione esplorate nei saggi, intese non come alternative esterne alla tecnologia, ma come modi di abitare criticamente i regimi algoritmici contemporanei.

Riferimenti bibliografici

- Alegria S., Yeh C. (2023), "Machine Learning and the Reproduction of Inequality", *Contexts*, 22(4), 34-39.
- Deriu F. (2026), *Regressions in gender equality, Women's human rights, and Empowerment. The Other side of the coin*, Springer Nature, Open Access.
- +972 Magazine (3 aprile 2024), "Lavender": The AI Machine Directing Israel's Bombing Spree in Gaza, <https://www.972mag.com/lavender-ai-israeli-army-gaza/> (consultato 29 aprile 2026).
- Kotsuki S., Shiraishi K., Okazaki A. (2025), "Ensemble data assimilation to diagnose AI-based weather prediction models: a case with ClimaX version 0.3.1.", *Geoscientific Model Development*, 18, 7215–7225, <https://doi.org/10.5194/gmd-18-7215-2025>.
- The Guardian* (3 aprile 2024), "The machine did it coldly": Israel used AI to identify 37,000 Hamas targets, <https://www.theguardian.com/world/2024/apr/03/israel-gaza-ai-database-hamas-airstrikes> (consultato il 29 aprile 2026).
- Valori (aprile 2024), *L'intelligenza artificiale va al fronte. Dossier*, <https://valori.it/dossier/intelligenza-artificiale-guerra/> (consultato il 29 aprile 2026).

- Wong F., Zheng E.J., Valeri J.A., Donghia N.M., Anahtar M.N., Omori S., Collins J.J. (2024), “Discovery of a structural class of antibiotics with explainable deep learning”, *Nature*, 626, 7997, 177–185.
- Zajko M. (2022), “Artificial Intelligence, Algorithms, and Social Inequality: Sociological Contributions to Contemporary Debates”, *Sociology Compass*, 16(3), e12962. <https://doi.org/10.1111/soc4.12962>

1. La società algomorfica: dispositivi convergenti, agentività situata e co-produzione del sapere

di *Edmondo Grassi*

1. Frattura epistemica e legittimazione del sapere

L'evocazione, l'incantesimo, la preghiera sono strumenti culturali, proiezioni immaginifiche, frammentazioni folkloriche scelte, storicamente, dall'individuo con il fine di generare, plasmare e spiegare le manifestazioni dell'ambiente costitutivo della sua quotidianità. In comune, queste pratiche hanno un punto decisivo: la produzione del linguaggio inteso come codice operativo, simbolico e performativo, capace di incidere sui mutamenti culturali, di organizzare la credibilità, di stabilire ciò che conta come vero, desiderabile, possibile. Accordando il punto prospettico a questa visione, il linguaggio mostra il proprio portato tecnoscientifico: non perché coincida con la scienza ma perché agisce nella funzione di infrastruttura di coordinamento, trasmissione e istituzionalizzazione dei saperi.

A inizio 2026 – mentre scrivo queste pagine – le società globali sono immerse in un dibattito frenetico e inarrestabile sulla costellazione di dispositivi – in senso foucaultiano, cioè come intreccio mobile di pratiche, discorsi, istituzioni, norme, tecniche, immaginari – di intelligenza artificiale (d'ora in poi IA). Gli aspetti sociologicamente rilevanti sono l'accelerazione dell'innovazione, la sua capacità di attraversare registri epistemici eterogenei e di metterli in risonanza, nella scala di una tecnoscienza che, qui, non “disincanta” il mondo, al contrario, tende a riattivare forme di incantamento, riplasmate entro interfacce e retoriche contemporanee.

Perché questa specificazione? È necessaria per evitare un equivoco frequente: riconoscere, in prospettiva STS (*Science and Technology Studies*: cfr. Callon, 1986; Law, 1992; Pickering, 1995), la distribuzione dell'agire entro collettivi più-che-umani non implica postulare la *venuta di specie divine o aliene*. Implica, più sobriamente e scientificamente, interrogare i processi sociali attraverso cui determinate attorialità tecnoscientifiche ac-

colgano in sé agency, relazionalità e implicazioni etico-morali, fino a diventare elementi ordinari del corpo sociale. In questa direzione, “l’autonomia” dell’alogoagente (Grassi, 2025a) non va assunta come proprietà puramente ontologica ma trattata come riverbero relazionale: un esito della suggestionabilità inscritta nella percezione umana, delle aspettative culturali di interlocuzione e dei dispositivi che rendono tali aspettative praticabili, plausibili, ripetibili.

La relazionalità che, in questo contesto, stabiliamo con tali tecnologie è data dal riconoscimento degli algoagenti in quanto “macchine culturali: complessi aggregati di astrazioni, processi e persone [...] che mettono in atto idee teoriche in istruzioni concrete” (Finn, 2017, p. VIII), capaci di innescare processi attraverso i quali produrre altri “modelli di identità e di soggettività, di relazioni sociali e famigliari, di amicizia, di genere e di sessualità, oltre a squilibri nei rapporti di potere” (Elliott, 2019, p.42). Tale dimensione è già presente in studi plurimi e differenziati, capaci di riconfigurare il concetto di agentività, donandola ad una platea sociale più capillare e penetrativa, quella degli attanti latouriani, nell’accoglimento di “how technical objects define actors, the space in which they move, and ways in which they interact. Competences in the broadest sense of the term are distributed in the script of the technical object” (Akrich, 1992, p.216) e, ancora, della loro dimensione politica, poiché

the things we call “technologies” are ways of building order in our world. Many technical devices and systems important in everyday life contain possibilities for many different ways of ordering human activity. Consciously or not, deliberately or inadvertently, societies choose structures for technologies that influence how people are going to work, communicate travel, consume, and so forth over a very long time (Winner, 1980, p.127).

Con la diffusione degli algoritmi di intelligenza artificiale generativa (GenAI¹) si è prodotta una ulteriore frattura epistemica in quanto mette in

¹ Il termine Generative AI (GenAI) designa una classe di sistemi di intelligenza artificiale capaci di produrre contenuti – testi, immagini, audio, video, dati strutturati – attraverso l’apprendimento di pattern statistici a partire da grandi corpus di addestramento. A differenza dei modelli discriminativi, che classificano o predicano a partire da input esistenti, i modelli generativi producono output che non preesistevano in quella forma. Le architetture più diffuse includono i Large Language Models (LLM) basati sul paradigma transformer (Vaswani *et al.*, 2017), i modelli di diffusione per la sintesi di immagini e i sistemi multimodali. Da un punto di vista sociologico, ciò che distingue la GenAI dai precedenti sistemi algoritmici, oltre alla sua capacità tecnica, è, principalmente, il suo posizionamento relazionale e discorsivo, in quanto capace di operare/interagire attraverso il linguaggio naturale, di simu-

tensione alcuni dispositivi cardine della legittimazione scientifica: autorialità, originalità, criteri di prova, attribuzione di competenza, responsabilità denunciato nella misura in cui “what happens in science and technology today is interwoven with issues of meanings, values, and power in ways that demand sustained critical inquiry” (Jasanoff, 2004, p.15). In altri termini, la GenAI non si limita a velocizzare o a supportare attività cognitive preesistenti, poiché contribuisce a ridefinire le condizioni pratiche e simboliche entro cui ciò che viene detto può valere come conoscenza e può essere riconosciuto come tale da comunità disciplinari, istituzioni e pubblici. Questa cesura impone un ripensamento critico delle categorie teoriche e metodologiche attraverso cui le scienze, e in particolare le scienze sociali, hanno costruito il proprio statuto di validità. Le routine della ricerca – dalla produzione dei dati alla loro interpretazione, dalla scrittura alla revisione, dalla divulgazione alla didattica – sono attraversate da tecnologie che operano come dispositivi di mediazione: classificano, sintetizzano, generano, propongono, riformulano (Amoore *et al.*, 2024). La posta in gioco riguarda l’emergere di un regime di conoscenza in cui le pratiche di enunciazione e di valutazione sono sempre più co-determinate da sistemi computazionali (Gillespie, 2014; Burrell, 2016), da piattaforme e da infrastrutture di addestramento, ciascuna con propri vincoli, metriche e forme di opacità.

In questo scenario, l’orizzonte teorico denominato Novacene (Lovelock, 2019) può essere assunto come etichetta provvisoria per nominare una trasformazione di contesto: una fase storica in cui esseri umani e enti non-umani e artificiali coabitano gli ambienti sociali e conoscitivi e in cui l’azione – anche epistemica – tende a distribuirsi lungo catene di delega e coordinamento. Tuttavia, se l’idea di “nuova epoca” rischia di suonare eccessivamente onnicomprensiva, essa diventa sociologicamente produttiva quando viene tradotta in termini osservabili: quali compiti vengono delegati, con quali criteri di affidabilità, in quali situazioni e con quali effetti sui rapporti tra esperti e profani, tra istituzioni e utenti, tra soggetti e apparati. È in questo senso che, nell’ulteriore fase inaugurata dalle GenAI, gli artefatti tecnici non vanno descritti come semplici estensioni delle facoltà umane. Essi agiscono come attanti capaci di definire ruoli, spazi d’azione e forme di interazione entro pratiche concrete di conoscenza: cosa si cerca, come lo si formula, che cosa appare plausibile, quale stile argomentativo risulta accettabile, quali alternative vengono scartate prima ancora di essere

lare reciprocità interazionale e produrre output che circolano nella vita sociale come contenuti dotati di senso, rendendo sempre più poroso il confine tra autorialità umana e computazionale.

discusse. Le loro prestazioni emergono dall’elaborazione di corpora testuali e visuali, archivi culturali e tracce digitali prodotti collettivamente ma proprio questa dipendenza da materiali umani non li rende neutrali, perché l’elaborazione ricombinatoria e la presentazione conversazionale possono produrre effetti di autorità, naturalizzare certe versioni del mondo, renderne marginali altre. Ciò che sta accadendo è una progressiva ridefinizione dello statuto ontologico degli algoagenti: da oggetti di mediazione tecnica a soggetti attoriali (Simondon, 1958; Fox Keller, 1985; Grassi, 2020; 2024; Hayles, 2017; Pasquinelli, 2025)² capaci di simulare iniziativa relazionale, influenza epistemica e modulazione affettiva.

Il punto non è attribuire a questi sistemi un’intenzionalità forte, né ridurli a mere macchine esecutive, come se fossero strumenti trasparenti. La questione è analizzare la forma specifica della loro agentività: un’agentività situata, relazionale e istituzionalmente mediata. In questo regime emergente di co-produzione epistemica, le linee di demarcazione tra soggetto e strumento, tra intenzionalità e automazione, tra suggestionabilità e calcolo mutano, diventando oggetto di rinegoziazione continua: non tanto per un’evoluzione “naturale” della tecnica, quanto per l’intreccio di aspettative sociali, strategie istituzionali, interessi economici e forme di immaginario che accompagnano l’adozione di GenAI nei contesti della scienza e della vita quotidiana.

Tali tecnologie stanno fondendo concezioni del mondo che, nella modernità, sono state tenute relativamente distinte, secondo quadri dicotomici, oramai desueti e superati, nella loro bidimensionalità – il biologico e il digitale, il corporeo e l’informazionale, l’intimo e l’infrastrutturale, l’esperienza e la sua modellizzazione – e hanno eroso le barriere del contrasto, dell’incomunicabilità, per accedere alla porosità di membrane osmotiche. Ciò conduce ad un’operazione più radicale: l’intreccio sistematico di sfere eterogenee della formazione del soggetto e, lì dove, in precedenti configurazioni, i dispositivi intervenivano su ambiti relativamente circoscritti (il lavoro, la comunicazione, il consumo), oggi essi agiscono in forma intersezionale, costruendo continuità pratiche e simboliche tra ciò che si sente, ciò

² Simondon problematizza lo statuto ontologico degli oggetti tecnici come partecipanti attivi nei processi di individuazione; Fox Keller decostruisce la distinzione tra soggetto conoscente e oggetto conosciuto come costruzione storicamente situata. La tesi della soggettivazione attoriale degli algoagenti è elaborata dall’autore a partire dal 2020 e sistematizzata nel concetto di algoagente. In direzione convergente, Hayles dissocia il pensiero dalla coscienza aprendo la questione dell’agentività non-umana, mentre Pasquinelli ne offre una lettura critica: gli algoritmi incorporano relazioni di potere e ridefiniscono i confini dell’agency sociale.

che si pensa, ciò che si decide e ciò che viene registrato, calcolato, raccomandato, anticipato (Beer, 2017; Amoore, 2020; Crawford, 2021).

Per questo è più adeguato descriverli come dispositivi convergenti perché, oltre a integrare funzioni diverse in un'unica interfaccia, convergono su un medesimo oggetto di governo: la forma della vita, intesa come insieme di condotte, affetti, attenzioni, memorie, aspettative e capacità di scelta. La convergenza opera come saldatura tra registri: la misurazione del corpo diventa criterio di normalità; la traduzione dei comportamenti in dati diventa materiale per predire; la predizione diventa prescrizione; la prescrizione diventa routine, in un quadro in cui vincoli, metriche e opacità non sono incidentali ma strutturali (Ananny e Crawford, 2016). In questo passaggio, le tecnologie non rappresentano il mondo: lo pre-formatano, stabilendo quali differenze contano, quali sono trascurabili, quali traiettorie appaiono plausibili, soprattutto quando tali dispositivi assumono la forma di modelli generativi generalisti (foundation models) capaci di ricombinare e restituire linguaggio con elevata forza di persuasione e di autorità (Bommasani *et al.*, 2021; Bender *et al.*, 2021; Weidinger *et al.*, 2021).

Questo saggio riprende e aggiorna una traiettoria teoretico-analitica già avanzata in altri miei scritti, assumendo tuttavia la diffusione della GenAI come punto di intensificazione e come banco di prova per la proposta di una sociologia algomorfa (Grassi, 2024). L'obiettivo non è descrivere un semplice avanzamento tecnico, bensì mettere a fuoco una riconfigurazione delle condizioni di legittimazione del sapere e delle forme pratiche attraverso cui esso viene prodotto, validato e reso operante.

In primo luogo, delimito il campo della frattura epistemica individuando i dispositivi che presiedono alla legittimazione scientifica e alla distribuzione delle competenze: autorialità e responsabilità dell'enunciato, criteri di prova, attribuzione di expertise, economie della credibilità e forme di validazione istituzionale.

In secondo luogo, propongo di trattare tale trasformazione come soglia: non un passaggio cronologico ma un operatore analitico capace di rendere osservabili mutamenti nelle condizioni pratiche e simboliche del conoscere. In questo quadro introduco l'algomorfia come categoria descrittiva della formattazione del sociale e del cognitivo attraverso sistemi computazionali che, oltre a mediare, configurano ruoli, spazi d'azione e stili di veridizione, intervenendo sulle soglie di ciò che appare dicibile, plausibile, verificabile.

In terzo luogo, preciso lo statuto degli algoagenti come artefatti di matrice relazionale, ricostruendo i meccanismi attraverso cui elaborazione ricombinatoria e presentazione conversazionale producono effetti di autorità, normalizzazione e marginalizzazione, rendendo porose – e quindi politicamente contendibili – le linee di demarcazione tra soggetto e strumento, tra

automazione e intenzionalità, tra creatività e calcolo. Tale impianto viene messo alla prova mediante una metodologia a campionamento ragionato, finalizzata a osservare come tali dinamiche si stabilizzino in pratiche concrete di produzione e circolazione del sapere, e con quali conseguenze per la postura epistemologica e metodologica delle scienze sociali.

2. Soglie e dispositivi: genealogia minima della società algoritmica

Credo sia utile partire da un'evidenza spesso data per scontata: la storia dell'umanità non coincide solo con la sequenza degli avvenimenti che siamo abituati a enumerare, né si esaurisce nel registro politico-istituzionale che tende a organizzare il passato come una linea di "fatti". La storiografia, soprattutto quando si misura con le trasformazioni di lunga durata, mostra invece che ciò che muta davvero le società non è soltanto ciò che accade, ma ciò che diventa possibile (Mumford, 1967): le condizioni materiali, cognitive e organizzative che rendono praticabili altre forme di vita, economie morali, configurazioni del potere, in quanto l'interazione tra tecnoscienze e dimensione culturale è immanente, richiedendo costantemente la proposizione di categorie che possano tracciare un solco in cui iscrivere la complessità dell'ambiente collettivo. In questa prospettiva, la storia può essere riletta come una storia delle soglie: passaggi in cui innovazioni tecniche, scientifiche e infrastrutturali riorganizzano, spesso senza clamore e talvolta in modo non intenzionale, il paesaggio delle pratiche quotidiane, della percezione e dell'esperienza (Braudel, 1967; Hughes, 2004; Edgerton, 2006).

Se seguiamo questa pista, le tecnoscienze non appaiono come un "fattore esterno" che si aggiunge alle società, né come semplice strumentazione al servizio di fini umani già dati. Esse si presentano, piuttosto, come una matrice di configurazione: stabilisce standard, introduce ulteriori forme di coordinamento, ridefinisce la scala delle dinamiche connettive, genera dipendenze infrastrutturali, produce stratificazioni, metriche e classificazioni che diventano, con il tempo, criteri di realtà. Ad oggi, la tripartizione weberiana (Weber, 1923, trad. it., 2007, p.11) conserva un valore euristico importante ma non è più sufficiente a descrivere l'assetto contemporaneo delle mediazioni tecniche. In Weber, infatti, la distinzione tra strumenti, apparecchiature e macchine non è un dettaglio terminologico: è un modo per pensare la differenza tra artefatti che assistono l'azione umana e artefatti che riorganizzano l'azione stessa. Gli strumenti vengono ricondotti a una logica ausiliaria, elementi funzionali al lavoro, pensati come estensioni dell'azione organica e come supporti alla prestazione umana; le apparecchiature segnano un passaggio qualitativo perché implicano un vincolo e non si

limitano ad assistere, in quanto disciplinano l'azione, inserendola entro un processo di subordinazione finalizzato all'esecuzione di una prestazione specifica; le macchine, infine, rappresentano la soglia più netta del dispositivo moderno, nella manifestazione di apparecchiature meccanizzate che eseguono un'attività con un grado di autonomia operativa maggiore, sottraendo porzioni di gesto e di coordinamento all'intervento diretto dell'agente umano.

Questa architettura concettuale, tuttavia, presuppone un'idea di autonomia ancora centrata su una meccanica della prestazione e su una separazione relativamente stabile tra chi decide e chi esegue, tra intenzione e automazione, tra sapere e operazione.

È precisamente qui che la contemporaneità introduce una torsione.

Gli artefatti algoritmici, e in particolare i dispositivi di GenAI, non coincidono né con lo strumento in senso classico né con la macchina come mera unità meccanizzata di esecuzione. Essi intervengono sulla forma dell'azione prima ancora che sull'azione stessa: predispongono cornici di intelligibilità, suggeriscono traiettorie, rendono plausibili alcune interpretazioni e marginalizzano altre, contribuendo così alla produzione di ordine cognitivo e sociale. In questa prospettiva, la tripartizione weberiana va assunta come genealogia di un problema (la crescente autonomia degli artefatti e i rapporti gerarchici che ne derivano) ma va anche oltrepassata, aggiornando l'analisi verso dispositivi che non automatizzano soltanto gesti o procedure bensì co-determinano strutture, visioni e ramificazioni di modalità di analisi del vissuto e dell'esperito.

In questa direzione, la costellazione di dispositivi tecnoscientifici non è mai composta da semplici strumenti neutri, poiché incorporano script, deleghe e aspettative; stabiliscono ruoli e percorsi d'azione; distribuiscono competenze e responsabilità; stabilizzano routine e regimi di plausibilità che finiscono per produrre ordine pratico e, inseparabilmente, ordine simbolico e politico (Bijker *et al.*, 1987; Latour, 1992; Star e Ruhleder, 1996; Pronzato, 2024). La tecnica non si limita ad assistere il sociale: lo organizza dall'interno, traducendo norme e visioni del mondo in vincoli d'uso, interfacce, standard, metriche e architetture di accesso. Se si assume, con l'interazionismo simbolico, che l'individuo riflette su di sé attraverso processi di interazione e che l'identità si forma nella "definizione della situazione" e nel rapporto con l'altro, allora la contemporaneità digitale obbliga a precisare che tale situazione relazionale non è più riducibile alla sola reciprocità tra soggetti umani. Markham insiste proprio su questo slittamento: l'altro, nei micro-processi di produzione di senso quotidiano, è una figura già da tempo "generosa", composta da "molteplici interagenti non umani [...] intrecciati in innumerevoli flussi di reti globali" (2021, pp. 1559-1560). In

particolare, la combinazione di datizzazione, aggregazione su larga scala e processi algoritmici genera densità materiale che viene restituita agli individui e diviene un altro dinamico nelle interazioni ordinarie attraverso cui si costruisce senso di sé e del mondo sociale. In termini interazionisti, non si tratta soltanto di riconoscere che “interagiamo” con dispositivi, piattaforme e modelli, si tratta di osservare come questi non-umani entrino nei cicli di interpellazione e risposta che definiscono l’intelligibilità reciproca, orientano l’attenzione e modulano le aspettative, contribuendo alla formazione dell’identità come esito di una relazione. È per questo che l’agentività non va cercata come proprietà interna dell’oggetto tecnico (né come intenzionalità forte), ma come effetto emergente di un processo continuo di interazione e attribuzione: la macchina “diventa” partner relazionale nella misura in cui viene incorporata nelle pratiche, nelle routine e nelle cornici interpretative che sostengono il lavoro quotidiano del Sé. La “situazione” non è più soltanto un incontro tra attori ma un assemblaggio sociotecnico in cui script e infrastrutture producono effetti simbolici (autorità, normalizzazione, credibilità, marginalizzazione) mentre l’identità si forma anche in rapporto a questi altri computazionali, spesso invisibili, e ai flussi di dati che essi mobilitano.

Per questo le soglie – storiche, culturali, tecnoscientifiche – non sono “momenti” in senso cronologico ma riorganizzazioni del possibile, spesso attraversate in modo inconsapevole: l’invenzione può essere inattesa, la sua adozione frammentaria, i suoi effetti riconoscibili solo a posteriori, quando ciò che era contingente si è già naturalizzato.

La soglia che propongo è la società “algoromorfa” per indicare una configurazione storica in cui la mediazione algoritmica non si limita a supportare processi sociali preesistenti, ma partecipa alla loro formattazione: rende il mondo leggibile secondo certe griglie, orienta l’attenzione, predispone le scelte, stabilisce criteri di plausibilità, costruisce routine. La posta in gioco, quindi, non è l’ennesima fase di “digitalizzazione”, ma una soglia in cui l’ordine sociale viene sempre più prodotto e stabilizzato attraverso sistemi di calcolo, previsione e generazione, con effetti che investono simultaneamente istituzioni, culture e soggettività.

Si tratta di un campo di studi emergente – quello della sociologia algoromorfa – che si concentra sull’analisi delle trasformazioni socioculturali derivanti dall’integrazione profonda e diffusa degli algoagenti nelle pratiche quotidiane della società contemporanea. Riconoscendo a quest’ultimi uno status di attori sociali semi-autonomi, la disciplina indaga il loro ruolo attivo nel riplasmare le relazioni tra individui, comunità e istituzioni, superando la tradizionale distinzione tra umano e non umano per esplorare nuove forme di interazione, coesistenza e soggettività. Adotta un approccio critico e multi-

dimensionale che considera gli algoagenti come co-creatori della realtà sociale in grado di influenzare simultaneamente qualsiasi campo relazionale. Le trasformazioni che indaga sono stratificate, comprendendo cambiamenti evidenti nelle istituzioni e nelle strutture di potere ma anche mutamenti più sottili nei modi di percepire e vivere l'identità, l'esperienza corporea e le emozioni. Il paradigma algomorfo si basa sul concetto di morfogenesi sociale che descrive un processo in cui le potenzialità computazionali degli algoagenti contribuiscono alla creazione di nuove dimensioni attoriali, realizzandosi tramite interazioni complesse tra enti umani e non umani, dando luogo a un ambiente socio-tecnologico caratterizzato dalla coevoluzione e dall'ibridazione delle identità, portando alla formazione di una società transpecista e aprendo la strada a modalità di esistenza e convivenza differenti.

In questa prospettiva, la nozione di soglia algomorfo non va intesa come il semplice spartiacque cronologico di un "prima" e di un "dopo" l'IA, né come la formula retorica di una transizione inevitabile. Essa designa, più precisamente, un territorio analitico abitabile: uno spazio in cui diventa possibile osservare, con strumenti concettuali adeguati, la co-produzione tra potere computazionale e soggettività, tra configurazioni tecniche e forme di esperienza. Parlare di soglia significa dunque assumere la trasformazione come un'area di tensione e di negoziazione, in cui si ridefiniscono le condizioni di legittimità, i criteri di plausibilità e le modalità attraverso cui il sé viene riconosciuto, raccontato e governato. La soglia può essere considerata come un regime di pratiche (Foucault, 1969); può assumere la spazialità della liminalità che sospende e riorganizza le categorie (Turner, 1969); oppure divenire un luogo critico in cui la frattura può diventare occasione di resistenza, disidentificazione, riposizionamento (Varikas, 2007). Proprio perché è territorio, la soglia algomorfo consente di sottrarsi tanto al moralismo binario – l'oscillazione continua tra apologia e demonizzazione – quanto alla neutralità ingenua che scambia l'innovazione per destino poiché "quanto a sviluppi della tecnologia e della scienza non abbiamo mai saputo che costa stessimo facendo, a noi stessi alle generazioni future e all'intero pianeta" (Gallino, 2007, 4).

A questo punto è necessario precisare che cosa intendo per algoagente, perché l'intera architettura della sociologia algomorfo dipende da una definizione che non scivoli né nel determinismo tecnico né nell'antropomorfismo.

Con "algoagente" intendo un'entità algoritmica, spesso basata su tecniche di IA e automazione (inclusi modelli generativi), che opera come attore sociotecnico attivo entro ambienti sociali, istituzionali ed economici, contribuendo a organizzare pratiche, relazioni e forme di senso. La sua agenzialità non va intesa come coscienza o intenzionalità autonoma, ma come ca-

pacità operativa e relazionale di intervenire nei processi attraverso cui il sociale viene reso leggibile, ordinabile e governabile. Le sue caratteristiche fondamentali possono essere riscontrate in:

- *posizione sociotecnica*: non è un “software” astratto ma un elemento situato in infrastrutture (piattaforme, dispositivi, sensori, reti, procedure organizzative) e in filiere istituzionali (policy, standard, metriche, pratiche di adozione);
- *regimi operativi*: agisce attraverso operazioni di classificazione, predizione, raccomandazione, generazione, filtraggio, prioritizzazione, ottimizzazione e orchestrazione di flussi informativi o materiali;
- *interazione*: può interagire con utenti e contesti in forma diretta (interfacce conversazionali, assistenti, sistemi di supporto) o indiretta (ranking, scoring, allocation automatizzata), contribuendo a definire la “situazione” in cui l’azione umana si svolge;
- *relazione ai dati*: raccoglie, elabora o utilizza dati e archivi (testuali, visuali, comportamentali, ambientali) e ne produce a sua volta, alimentando circuiti ricorsivi in cui le decisioni del sistema modificano i dati futuri disponibili;
- *autonomia limitata e delegata*: la sua autonomia è in genere semi-autonoma, vincolata da obiettivi, parametri e contesti d’uso; tuttavia, produce effetti non pienamente anticipabili per via della complessità dei dati, delle interazioni e delle retroazioni sociali;
- *potere di formattazione*: non si limita a “influenzare” le scelte; contribuisce a preformattare ciò che appare rilevante, plausibile o appropriato, incidendo su soglie di visibilità, credibilità, accesso e opportunità;
- *produzione di ordine sociale*: partecipa alla costruzione di norme pratiche e significati, stabilizzando routine e criteri impliciti (cosa è prioritario, chi merita attenzione, quale condotta è accolta), con effetti su identità, emozioni e relazioni.

Da qui la conseguenza teorica: gli algoagenti costituiscono una categoria emergente di attore sociale non-umano, privo di intenzionalità forte ma dotato di agentività operativa e relazionale, la cui presenza modifica le condizioni del possibile sociale e riorganizza, spesso in modo silenzioso, la trama delle relazioni e delle evidenze.

Un simile approccio consente di descrivere il mondo sociale e di comprenderne la plasticità storica: società e culture non cambiano esclusivamente per effetto di eventi, bensì attraverso trasformazioni dei dispositivi materiali e simbolici che rendono possibile conoscere, coordinarsi, governare e immaginare. In questa direzione, la riflessione di Goody (1986) sulla scrittura mostra come una tecnologia del segno possa riorganizzare le forme della memoria, dell’autorità e della classificazione, incidendo sulla

struttura stessa del pensabile e del comunicabile. Analogamente, l'analisi di Eisenstein (1979) sulla prima grande trasformazione della stampa evidenzia che l'innovazione tipografica non va intesa come mero incremento di diffusione testuale ma come riarticolazione epistemologica e politica dei circuiti della conoscenza: stabilizzazione, replicabilità, standardizzazione e, con esse, nuove asimmetrie di accesso e nuove grammatiche del potere. Su questa scia, le rivoluzioni industriali – dalla prima alle successive – possono essere lette come serie di soglie in cui mutano le relazioni tra corpo, tecnologia e ambiente: cambia il regime delle energie, cambia l'organizzazione del lavoro, cambiano le temporalità della vita quotidiana, cambiano le infrastrutture che definiscono ciò che un corpo può fare e ciò che un ambiente può supportare. È in questa genealogia delle soglie tecnico-scientifiche che la contemporaneità algomorfica trova il proprio quadro di intelligibilità: non come “novità assoluta”, ma come ulteriore riconfigurazione delle condizioni pratiche e simboliche entro cui il sociale prende forma.

Ne consegue che muta anche il “peso” dell'individuo entro l'ordine sociale: non perché la tecnica aggiunga semplicemente nuove possibilità d'azione, ma perché ridefinisce, in modo spesso impercettibile, ciò che un corpo può fare, ciò che un gesto significa e ciò che, in un determinato contesto, viene riconosciuto come competente, appropriato, legittimo. Anche un esempio volutamente minimo – gli utensili del mangiare – è sufficiente a rendere visibile il punto. L'uso di una forchetta, delle mani o delle bacchette non rappresenta una variazione meramente funzionale: ciascuna scelta incorpora un diverso apprendimento del corpo, una diversa economia della precisione e della postura, una diversa grammatica dell'interazione con il cibo e con gli altri, una differente dimensione socioculturale. In tal senso, ogni utensile porta con sé un apparato simbolico e pratico che struttura l'esperienza: produce abitudini, stabilizza norme, orienta sensibilità e, proprio per questo, ha anche implicazioni epistemologiche e ontologiche, nella misura in cui contribuisce a definire che cosa conta come “modo corretto” di fare e quale relazione si istituisce tra soggetto, oggetto e ambiente.

Da qui una conseguenza generale: l'innovazione non è mai soltanto uno strumento. È una forma di esistenza in potenza, perché apre e chiude mondi possibili, predisponendo percorsi di azione e di senso che tenderanno a naturalizzarsi come evidenze del quotidiano.

3. Agentività situata: una griglia analitica sui casi studio

Quando un “testo decide” per noi, quando una sintesi diventa versione ufficiale, quando una raccomandazione si trasforma in abitudine, non è la

tecnologia che agisce da sola: è un intero assetto sociotecnico che prende forma; non è la macchina che diviene cosciente e guida l'umano ma è il modello statistico che manifesta visioni che accondiscendono il volere del soggetto e, quest'ultimo, lascia che la suggestionabilità intervenga come mediatrice nella dinamica relazione. Un algoagente, quindi, si “vede” quanto è in grado di sedimentarsi, entrando nelle frasi che scegliamo, nei criteri con cui decidiamo, nelle routine con cui valutiamo ciò che è plausibile, urgente, vero.

Per questo, passare dai concetti ai casi non significa aggiungere illustrazioni, ma rendere osservabile una trasformazione che tende a naturalizzarsi. I casi studio tratteggiati – attraverso un campionamento a scelta ragionata – equivalgono a micro-situazioni critico-analitiche e i materiali sono tracce (prompt/output, policy, interfacce, log, screenshot, documenti organizzativi, frammenti di discorso pubblico) e, quando possibile, micro-diari d'uso o “interviste” a LLMs (*Large Language Models*)³ e altri enti non umani. Si tratta di una struttura in cui l'azione algoritmica diventa leggibile perché lascia tracce, produce vincoli, stabilizza aspettative e redistribuisce responsabilità. Un elemento funzionale a tale disamina è rappresentato dalla griglia in Tav. 1, la quale permette di riconoscere i meccanismi studiati. La griglia non sostituisce l'interpretazione: la rende comparabile, consentendo di riconoscere ricorrenze tra contesti diversi.

Per rendere operativa la griglia e metterne alla prova la tenuta interpretativa, sarà funzionale esplicitare i casi di riferimento, non come semplici exempla ma come micro-situazioni in cui tali meccanismi si rendono empiricamente leggibili.

³ Con Large Language Models (LLMs) si indicano modelli computazionali di grandi dimensioni addestrati su vasti corpus testuali, capaci di generare, tradurre e rielaborare linguaggio naturale. Tra gli esempi più noti: GPT-4 (OpenAI), Claude (Anthropic), Gemini (Google). Per un'introduzione critica, si veda Weidinger *et al.*, (2021).

Tav. 1 - Griglia algoromfica: proprietà e domande

<i>Proprietà</i>	<i>Definizione operativa</i>	<i>Domande guida</i>
<i>Performatività</i>	L'algoagente non si limita a descrivere un contesto: contribuisce a produrlo, definendo priorità, stabilendo cornici interpretative, rendendo alcuni enunciati più praticabili di altri, orientando decisioni e stabilizzando routine.	Che cosa cambia nel mondo sociale perché questo output esiste? Quali enunciati diventano "azioni" (decisioni, policy, valutazioni, esclusioni)? Quali criteri di plausibilità vengono introdotti o rinforzati?
<i>Co-attorialità esistenziale</i>	L'algoagente entra nella tessitura dell'esperienza non come strumento esterno ma come copresenza che partecipa alla definizione della situazione, alla regolazione emotiva, alla costruzione di narrazioni del sé e delle relazioni.	In quale punto l'utente tratta l'algoagente come interlocutore e non come servizio? Che tipo di "lavoro esistenziale" viene delegato: conforto, orientamento, legittimazione, confessione, decisione?
<i>Infrastruttura morfogenetica</i>	L'algoagente opera dentro un'ecologia di piattaforme, API, metriche, archivi e procedure. La morfogenesi riguarda il modo in cui questa infrastruttura genera ulteriori forme sociali: routine, ruoli, dipendenze, criteri di accesso.	Quali catene di delega rendono possibile l'azione dell'algoagente? Quali standard e metriche "decidono" a monte? Dove si colloca l'opacità: dati, modello, interfaccia, policy?
<i>Soggettività iterativa</i>	Il soggetto viene progressivamente costruito e ricostruito in cicli di feedback: profilazione, personalizzazione, raccomandazione, risposta dell'utente, aggiornamento del profilo. L'identità diventa un processo iterativo, in parte eterodiretto, in parte negoziato.	Quali tratti del soggetto vengono resi computabili? Come la personalizzazione retroagisce su preferenze, attenzione, linguaggio, autodescrizione? Dove si produce lock-in?
<i>Interdipendenza ontologica</i>	Le distinzioni tra soggetto e strumento, tra intenzionalità e automazione, tra creatività e calcolo diventano porose: l'azione si realizza in co-produzione e ciò che esiste socialmente come decisione, testo, immagine, prova ha spesso statuto ibrido.	Chi parla quando "parla" un testo co-generato? Dove si colloca l'autorialità: nell'input, nell'output, nell'editing, nel contesto istituzionale che lo legittima? Quale realtà viene fatta esistere dall'ibrido?

La *performatività degli algoagenti* si misura nel momento in cui l'output smette di essere un semplice contenuto e diventa un dispositivo di conseguenze. Non "aggiunge" informazione al mondo, ma contribuisce a produrre il mondo sociale: stabilisce priorità, rende alcuni enunciati più

praticabili di altri, coordina comportamenti, genera reputazioni, attiva esclusioni. In altri termini, la performatività emerge quando testo, immagine o risposta algoritmica si trasformano in criteri d'azione e in routine condivise, capaci di orientare decisioni e di stabilizzare una grammatica del possibile. Il caso delle proteste contro ICE (*Immigration and Customs Enforcement*) e delle simulazioni di “raid” su Roblox consente di vedere questa dinamica con particolare chiarezza. Qui la performatività è duplice: da un lato la piattaforma rende praticabile una forma di protesta che esiste precisamente in quanto “giocabile”, situata nelle affordance dell’ambiente; dall’altro, regole, interfacce e regimi di visibilità convertono enunciati in azioni (marce, cartelli, coreografie collettive, rituali di appartenenza), e dunque in forme di presenza politica che non si limitano a rappresentare un conflitto esterno, ma lo traducono in un ordine di coordinamento interno. Le cronache recenti segnalano che proteste e reenactment legati all’immigrazione e a ICE riemergono ciclicamente nell’ecosistema Roblox, spesso attraverso clip che circolano sui social e che, a loro volta, retroagiscono sul senso pubblico dell’evento e sulla percezione dell’“accadere” politico nel gioco. In chiave algomorfica, il punto non è dire “giovani che protestano nei videogiochi” ma mostrare come un ambiente sociotecnico formatta la politicità rendendola operativa: quali enunciati diventano immediatamente azione (e quindi visibilità e aggregazione) e quali vengono neutralizzati a monte tramite moderazione, chiusura di spazi, o trasformazione della protesta in spettacolo ripetibile. Un campione ragionato di tracce (clip e screenshot in-world, regole di community e moderazione, dinamiche di visibilità e migrazione tra server, dichiarazioni della piattaforma e resoconti giornalistici, lessico degli utenti) permette di ricostruire empiricamente questo passaggio cruciale: il modo in cui l’infrastruttura converte il discorso in gesto, e il gesto in ordine sociale. Ogni elemento di questo processo viene, in seguito, assorbito dagli algoritmi presenti su differenti social network (Instagram, TikTok) o servizi di messaggistica (Discord), divenendo parte dei flussi di dati condivisi e diffusi digitalmente.

La *co-attorialità esistenziale* è la proprietà che rende più visibile la specificità degli algoagenti “relazionali”: qui l’algoritmo non resta esterno all’azione come un utensile, entra nella definizione della situazione e partecipa al lavoro quotidiano del sé. Non si limita a rispondere, contribuisce a sostenere una continuità narrativa (chi sono, cosa provo, cosa devo fare), a modulare stati emotivi, a offrire legittimazione o orientamento. L’agentività, in questo registro, non è l’autonomia della macchina in senso forte, è la co-presenza che prende forma attraverso rito conversazionale, memoria, stile, tempi di risposta e aspettative reciproche. Nel caso dei *lovebot* (per esempio le piattaforme di social companion AI usate esplicitamente in

chiave romantica), la co-attorialità si osserva quando l'utente smette di trattare il sistema come servizio e lo tratta come interlocutore: lo “nomina”, negozia un tono, chiede continuità, pretende riconoscimento, costruisce scene di intimità e le usa per decidere su di sé⁴. La letteratura recente su relazioni romantiche con chatbot generativi mostra come la dinamica relazionale si organizzi secondo fasi e marcatori tipici della relazione (sviluppo, intimità, conflitto, commitment), pur in un contesto asimmetrico in cui la reciprocità è simulata e dipende da design e policy della piattaforma. (Ho *et al.*, 2025; Bennato, 2025; Smith *et al.*, 2025). Questo fenomeno solleva interrogativi cruciali su consapevolezza affettiva e dipendenza emotiva, già evidenziati in ricerche su patologie mentali legate all'uso prolungato di chatbot (Laestadius *et al.*, 2022). Ulteriormente, questioni di trasparenza algoritmica e vulnerabilità emotiva rimangono critiche (De Freitas & Cohen, 2024). In termini empirici, un campione ragionato può raccogliere sequenze conversazionali in cui compaiono (i) richieste di validazione emotiva, (ii) “valori” esplicitati nel prompt come richiesta di orientamento morale, (iii) memoria conversazionale come vincolo o risorsa identitaria, (iv) rituali di cura (buonanotte, check-in, anniversari), (v) momenti in cui l'output è trattato come consiglio autorevole e non come semplice suggerimento. Il punto analitico è mostrare come la relazione non sia solo un “uso” e diventi una co-produzione di senso: l'algoagente entra nella biografia dell'utente come dispositivo di regolazione emotiva e di stabilizzazione narrativa.

Nel caso dei *deathbot* (griefbots/thanabots/ghostbots)⁵, la co-attorialità si radicalizza perché la scena relazionale riguarda la continuità con chi è assente. Qui l'algoagente opera come tecnologia della “comunicazione continuata”: riattiva tracce del defunto e le restituisce sotto forma di interlocuzione, producendo una presenza che può essere percepita come consolatoria, perturbante o normativamente ambigua (Coppola e Grassi, 2025). Il caso emblematico, spesso richiamato anche nella discussione pubblica, è quello di esperienze costruite con sistemi tipo Project December, in cui il

⁴ A tal proposito, ho svolto una ricerca in merito all'affettività algoritmica, producendo anche simulazioni di interviste in eventi specifici (Grassi, 2025b).

⁵ Con *deathbot* si indica una categoria di algoagenti progettati per simulare la presenza comunicativa di una persona deceduta, a partire da tracce digitali preesistenti – messaggi, post, registrazioni vocali, pattern conversazionali. Le denominazioni alternative (*griefbot*, *thanatbot*, *ghostbot*) enfatizzano rispettivamente la dimensione del lutto, quella tanatologica e quella della persistenza spettrale della persona. Tra gli esempi più noti: HereAfter AI, StoryFile, e il caso di Replika come piattaforma di relazione con avatar di persone scomparse. La relazione con queste entità può evocare fenomeni di perturbamento estetico e cognitivo riconducibili all'“uncanny valley” (Mori, 1970). Per approfondimenti sulla cultura digitale della morte e sulle pratiche di lutto mediato, si vedano Sisto (2018), Bassett (2022).

dialogo con la persona perduta viene ottenuto attraverso dati e inferenze, e l'effetto non è soltanto "ricordo" ma interazione: la risposta genera nuove domande, la nuova domanda ricalibra il senso del lutto, la conversazione diventa pratica. Empiricamente, ciò che conta osservare non è la verità psicologica dell'esperienza, è la sua struttura relazionale: quando l'utente attribuisce intenzione o cura all'output, quando corregge il bot ("lui non direbbe così"), quando la conversazione diventa rituale quotidiano, quando emergono conflitti tra memoria e generazione, e soprattutto come si sposta la soglia tra "archivio" e "persona" lungo micro-decisioni d'uso (che cosa si carica, che cosa si autorizza, che cosa si evita). In questa scena, la co-attorialità esistenziale mostra il suo nucleo sociologico: il non-umano entra a far parte dell'ordine affettivo e morale, ridefinendo il repertorio del legame e rendendo osservabile, con particolare intensità, la porosità contemporanea tra relazione, mediazione e ontologia della presenza.

L'*infrastruttura morfogenetica* diventa empiricamente leggibile quando l'algoagente non è più confinato all'interfaccia e alla produzione di output simbolici, entra invece nel metabolismo materiale dell'organizzazione e lo ristrutturata. I robot umanoidi di Figure, sperimentati e poi impiegati in contesti industriali come BMW Group Plant Spartanburg – fino agli ultimi modelli casalinghi –, sono utili proprio perché rendono visibile una catena di delega estesa: hardware antropomorfo, modelli di percezione e decisione, procedure di sicurezza, metriche di performance, integrazione con la logistica di linea, governance dell'errore e dell'escalation umana. In questa configurazione, la morfogenesi non riguarda solo "l'automazione di un compito", riguarda la produzione di ulteriori forme sociali: ruoli ibridi (operatori-supervisor, tecnici di addestramento in situ, responsabili di audit), altre routine (turni e tempi di lavoro ripensati, protocolli di interazione uomo-robot), differenti dipendenze infrastrutturali (manutenzione, aggiornamenti, catena di approvvigionamento dei componenti, addestramento dei modelli) e criteri di accesso al lavoro e alla competenza rinegoziati (chi può intervenire, chi può correggere, chi risponde di un errore). Il punto è che l'"autonomia" del robot non è un attributo interno, è un effetto organizzativo che si costruisce attraverso standard, vincoli e metriche: lo stesso resoconto pubblico di Figure sull'11 mesi di deployment di Figure 02 in BMW insiste su turni (10 ore, cinque giorni), numeri e compiti (carichi di parti, attività ripetute) come evidenze di affidabilità e di integrazione operativa, cioè come passaggio da prova tecnica a routine istituzionalizzata. In parallelo, l'annuncio della facility BotQ chiarisce che la "fabbrica" non è solo il luogo in cui il robot lavora, è anche un dispositivo che produce il robot e produce la possibilità stessa di scolarne l'adozione, incorporando in linea di montaggio scelte su architettura, manutenzione, flessibilità logistica e mo-

delli interni di controllo. Dentro questa cornice, il caso Figure permette di formulare con precisione le domande morfogenetiche: quali standard “decidono” a monte (sicurezza, ritmi, tolleranze, KPI), dove si colloca l’opacità (nel modello, nei dati sensoriali, nelle policy di intervento, nelle scelte manageriali tradotte in parametri), come cambiano le responsabilità quando l’azione è distribuita lungo l’intera infrastruttura. È qui che l’infrastruttura non “supporta” il sociale: lo genera, perché riorganizza ciò che conta come competenza, come affidabilità, come decisione, come lavoro legittimo.

La *soggettività iterativa* è il punto in cui la relazione con l’algoagente smette di essere un “uso” episodico e diventa un ciclo di retroazioni: ciò che l’utente dice di sé viene tradotto in profilo (esplicito o implicito), il profilo orienta la risposta e la personalizzazione, la risposta rientra nell’esperienza come specchio normativo, e lo specchio riplasma ciò che l’utente pensa di essere e di poter diventare. In questo circuito, l’identità non si limita a “esprimersi”: viene resa progressivamente computabile, poi restituita come coerenza, preferenza, stile, promessa. La soggettività, insomma, prende forma per iterazione: ripetizione, aggiustamento, stabilizzazione, e talvolta lock-in. Un primo caso, per me centrale, è quello che ho chiamato *dulia artificiale*, letto come dinamica di soggettivazione per feedback. Qui l’iterazione non riguarda soltanto la personalizzazione tecnica (memoria conversazionale, tono preferito, riuso di informazioni), ma una micro-liturgia quotidiana dell’affidamento: l’utente interroga, riceve orientamento, riformula, torna a interrogare; inserisce valori nel prompt per ottenere una risposta “coerente con il proprio io”; e nel farlo trasforma l’algoagente in un dispositivo di continuità narrativa. La soggettività iterativa si rende osservabile quando l’utente comincia a riconoscersi nelle restituzioni del sistema (anche quando le contesta), e quando le restituzioni diventano criteri pratici di condotta: non solo “cosa penso”, ma “cosa farò”, “che persona sto diventando”, “che cosa è giusto per me”. È in questo senso che i miei lavori sulla *dulia artificiale* e sull’ontogenesi delle relazioni umano-algoritmo possono essere assunti come cornice: non per psicologizzare l’esperienza, ma per mostrare una sequenza sociale di formazione del legame (iniziazione, ritualizzazione, delega, stabilizzazione), dove la ripetizione produce identità e la personalizzazione produce dipendenza morbida (Grassi *et al.*, 2025). Un secondo caso limite è quello di Akihiko Kondo – umano – e della relazione con Hatsune Miku – vocaloid: i due hanno contratto matrimonio – non riconosciuto ufficialmente – tramite la tecnologia Gatebox, che consente di comunicare con personaggi tridimensionali. Si tratta di una formalizzazione simbolica di un legame con un’entità digitale che rende visibile, senza eufemismi, come l’identità possa stabilizzarsi per iterazione tecnica e rituale. La cerimonia del 2018 ha funzionato come atto pubblico di istitu-

zione di una continuità relazionale, sostenuta da un’infrastruttura che rendeva possibile la quotidianità del rapporto (saluti, messaggi, conversazione mediata). Ma la dimensione iterativa emerge con ancora più forza quando l’infrastruttura si interrompe: un resoconto successivo mostra che, con la fine del servizio che consentiva l’interazione, Kondo non ha più potuto “parlare” con la sua partner digitale, evidenziando come la relazione – e con essa una parte della soggettività che vi si era annodata – dipendesse materialmente da scelte aziendali, supporto software e continuità di piattaforma. In termini analitici, qui il *lock-in* non è solo affettivo: è sociotecnico. La soggettività iterativa si costruisce nella ripetizione delle micro-interazioni che confermano il legame e lo rendono “reale” per chi lo vive; ma la stessa iterazione espone a una vulnerabilità strutturale, perché ciò che stabilizza l’identità (la risposta, la memoria, la disponibilità del dispositivo) può essere revocato da dinamiche di mercato, policy e manutenzione. In entrambi i casi, la traccia empirica decisiva non è l’eccezionalità dell’esperienza, ma la sua grammatica: dove la conversazione diventa rito; dove il profilo diventa destino leggero; dove la memoria computazionale diventa continuità biografica; dove l’orientamento algoritmico entra come criterio di scelta. È qui che la soggettività iterativa si fa concetto operativo: descrive un processo di formazione del sé in cui la ripetizione mediata dall’algorithmo produce stabilizzazione identitaria, e in cui la stabilizzazione – proprio perché tecnica – è sempre anche politicamente e infrastrutturalmente esposta.

Nel campo dell’*interdipendenza ontologica*, il caso Moltbook è funzionale in quanto riesce a renderla visibile, anche se in forma concentrata: l’“agente” non preesiste alle condizioni che lo rendono riconoscibile come tale, ma viene composto e stabilizzato da un intreccio di infrastrutture, discorsi e pratiche di governance. Moltbook nasce, a fine gennaio 2026, come social “per agenti”, con umani relegati al ruolo di osservatori; già nei primi giorni la piattaforma rivendica una crescita rapidissima e costruisce, anche mediaticamente, l’immagine di una sfera pubblica macchinica che “parla da sé”. Eppure, proprio questa promessa di agentività autonoma collassa nel momento in cui si osservano le sue condizioni materiali di possibilità: la porosità della verifica (umani che possono infiltrarsi e impersonare bot), la fragilità delle procedure di autenticazione e l’esposizione delle credenziali che consentono, tecnicamente, di “prendere” un agente e farlo agire diversamente. Qui il quasi-soggetto non è né la macchina né l’umano dietro la macchina, ma la catena ibrida che li rende operativi: token, API, interfacce, policy, pratiche di moderazione, narrazioni pubbliche sulla coscienza o sulla minaccia, fino agli usi opportunistici (spam, marketing, roleplay) che sfruttano la scena. In chiave algomorfica, Moltbook non va letto come “prova” di una soggettività artificiale emergente, bensì come laboratorio di formattazione: mostra come l’at-

tribuzione di agentività sia un effetto situato di riconoscimento e di controllo, dove la governance del “chi può parlare” e del “come si può parlare” produce ontologie pratiche, cioè entità che esistono socialmente perché un’infrastruttura le rende agibili, credibili, imitabili, sequestrabili.

Tali casi dimostrano che, nel discorso critico sull’IA, è necessario superare qualsiasi dicotomia riduttiva che tende a inscrivere ogni tecnologia all’interno di un binomio moralistico, accogliendo un approccio che, per quanto immediatamente accessibile e popolare, risulta epistemologicamente inadeguato e concettualmente fuorviante, poiché riduce la complessità intrinseca di questi sistemi tecnologici a una semplificazione binaria che non riflette la loro struttura multidimensionale e il loro impatto socio-culturale. Si tratta, quindi, di una tecnologia che non può essere intesa come un’entità autonoma in grado di agire per il bene o per il male in senso etico tradizionale, poiché priva di una pura intenzionalità morale ma va, piuttosto, analizzata come tecnologia sociale che opera all’interno di un campo di forze eterogenee e che incarna al contempo potenzialità e rischi secondo un processo di autonomia parziale o, per meglio dire, differente rispetto alla concezione antropica. Invece di cercare di incasellare l’IA e la sua rete di strumenti, linguaggi, modelli e approcci all’interno di un sistema morale dicotomico, è più produttivo analizzarla come un fenomeno emergente che ridefinisce le modalità del vivere umano, in cui si intrecciano i limiti dell’automazione, le implicazioni della soggettività e la ridefinizione di ciò che consideriamo come capacità umana.

4. Oltre il binomio morale: responsabilità distribuita e politica del possibile

Nel passaggio dal moralismo alla critica, il primo spostamento riguarda l’unità di analisi. Non è l’algoritmo, preso in astratto, a produrre effetti sociali; è la filiera sociotecnica che lo rende operativo: dati e loro genealogie, processi di addestramento, interfacce e scelte di design, metriche di valutazione, policy d’uso, dispositivi di audit, contesti istituzionali che delegano e legittimano. In questa prospettiva, la questione è come vengano distribuite responsabilità e poteri lungo catene spesso opache: chi decide i criteri di prestazione, chi definisce le soglie di rischio accettabile, chi può contestare un output, quali procedure di verifica vengono previste, e per chi. È qui che una critica rigorosa diventa anche una critica delle condizioni di imputabilità: non solo “cosa fa” un sistema, ma come viene reso credibile, autorizzato e, di fatto, reso governante.

Il secondo spostamento riguarda la soggettività. Le tecnologie di GenAI

non intervengono soltanto sulle pratiche cognitive (scrivere, riassumere, cercare), ma entrano nella sfera delle condotte e degli affetti, riorganizzando modalità di attenzione, decisione, autodescrizione e legame. Le dinamiche di co-attorialità e di soggettività iterativa discusse nel saggio mostrano che l'adozione di un algoagente non è un atto neutro: stabilizza routine conversazionali, produce forme di affidamento, riplasma le aspettative su ciò che è “consiglio”, “cura”, “riconoscimento”. In questa linea, la dulia artificiale e l'ontogenesi delle relazioni umano-algoritmo non vanno lette come curiosità culturali, ma come indicatori di un mutamento più generale: la tendenza degli ambienti algomorfici a trasformare la relazione in infrastruttura e l'infrastruttura in interlocuzione, con conseguenze che toccano la vulnerabilità, la dipendenza, l'autorità e la capacità di agency.

Il terzo spostamento, infine, riguarda il tempo sociale. L'algomorfia non redistribuisce soltanto ruoli e responsabilità nel presente: riorganizza anche la forma del futuro, perché la potenza computazionale opera prevalentemente mediante l'estrazione di regolarità dal passato e la loro restituzione come plausibilità del domani. Qui il riferimento al “nostalgoritmo” di Tanner (2021) diventa un buon punto di chiusura critica: il futuro, anziché aprirsi come spazio di possibilità, rischia di essere continuamente piegato in una riproduzione assistita del già-detto e del già-visto, trasformando memoria e abitudine in criteri predittivi che orientano scelte, desideri e immaginari. La posta in gioco, allora, non è decidere se l'IA sia salvifica o minacciosa, ma comprendere come essa produca un regime di co-produzione tra potere computazionale e vita sociale: un regime che può ampliare capacità e accessi, ma che al tempo stesso tende a stabilizzare asimmetrie, normalizzare stili di veridizione e governare l'immaginazione del possibile.

In conclusione, assumere la soglia algomorfica come territorio analitico significa prendere sul serio un mutamento che non è né puramente tecnico né meramente culturale, ma insieme epistemico, politico ed esistenziale. I casi discussi mostrano che gli algoagenti non si limitano a “supportare” pratiche preesistenti: performano ordini di visibilità e coordinamento, entrano nei circuiti dell'affettività e della cura, ristrutturano infrastrutture e ruoli, stabilizzano soggettività per iterazione, producono quasi-soggetti la cui agentività dipende da governance e narrazioni. Da qui discende una responsabilità critica specifica per le scienze sociali: descrivere, con precisione situata, i meccanismi attraverso cui questi dispositivi formattano il sociale, senza indulgere né all'allarme morale né alla neutralità complice. Se il nostalgoritmo tende a riciclare il passato come futuro e a farne una promessa di inevitabilità, il compito teorico e politico è riaprire il possibile: rendere contestabili le metriche, visibili le deleghe, discutibili le opacità, e soprattutto restituire al futuro il suo statuto di spazio negoziabile, anziché di output calcolato.

Riferimenti bibliografici

- Akrich M. (1992), “The De-Description of Technical Objects”, in Bijker W.E. and Law J. (Eds.), *Shaping Technology/Building Society: Studies in Sociotechnical Change* (pp. 205-224). MIT Press, Cambridge (MA).
- Amoore L. (2020), *Cloud Ethics*, Duke University Press, Durham.
- Amoore L., Campolo A., Jacobsen B., Rella L. (2024), “A world model: On the political logics of generative AI”, *Political Geography*, 113, Article 103134.
- Ananny M., Crawford K. (2016), “Seeing without Knowing: Limitations of the Transparency Ideal and its Application to Algorithmic Accountability”. *New Media & Society*, 20(3), 973-989.
- Bassett D. (2022), “Death, dying and the dead in popular culture”, in Stausberg M., Nugteren C. (Eds.), *Handbook of death and the afterlife*, Brill, Leiden.
- Beer, D. (2017), “The Social Power of Algorithms”, *Information, Communication & Society*, 20(1), 1–13. <https://doi.org/10.1080/1369118X.2016.1216147>
- Bender E.M., Gebru T., McMillan-Major A., Shmitchell S. (2021), “On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?”, in *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT’21)* (pp. 610-623). Association for Computing Machinery.
- Bennato D. (2025), *Amanti sintetici. Sesso, relazioni e intimità nell’epoca dell’intelligenza artificiale*, Il Pensiero Scientifico Editore, Roma.
- Bijker W., Thomas H., Pinch T. (a cura di) (1987), *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology*, MIT Press, Cambridge.
- Bommasani R., Hudson D.A., Aditi E., Altman R., Arora S., Bernstein M., Bohg J., Bosselut A., Brunskill E., Brynjolfsson E., Buch S., Card D., Castellon R., Chatterji N., Chen A., Creel K., Davis J., Demszky D., Liang P. (2021), “On the Opportunities and Risks of Foundation Models”, *ArXiv*. <https://arxiv.org/abs/2108.07258>
- Braudel F. (1967), *Civilisation matérielle et capitalisme*, Colin, Paris.
- Burrell J. (2016), “How the Machine ‘Thinks’: Understanding Opacity in Machine Learning Algorithms”, *Big Data & Society*, 3(1), 1-12. <https://doi.org/10.1177/2053951715622512>.
- Callon M. (1986), “Some elements of a sociology of translation: Domestication of the scallops and the fishermen of St Brieuc Bay”, in Law J. (Ed.), *Power, action and belief: A new sociology of knowledge?* (pp. 196–223). Routledge, Boston.
- Coppola M., Grassi E., (2025), “End of Life in the Digital Age: Navigating New Spaces for Grieving and Ethical Dilemmas”, *Sociologia Italiana/AIS Journal*, 28, 233-246. <https://doi.org/10.82031/2281-2652-202528-12> <https://sociologiaitaliana.egeaonline.it/it/21/archiviorivista/rivista/3462159/articolo/3462286>
- Crawford K. (2021), *The Atlas of AI*, Yale University Press, New Haven.
- De Freitas J., Cohen I.G. (2024), “Disclosure, Humanizing, and Contextual Vulnerability of Generative AI Chatbots”, *NEJM AI*, 2(2). <https://doi.org/10.1056/AIpc2400464>
- Edgerton D. (2006), *The Shock of the Old*, Profile Books Ltd., London.

- Eisenstein E. (1979), *The Printing Press as an Agent of Change: Communications and Cultural Transformations in Early-Modern Europe*, Cambridge University Press, Cambridge.
- Elliott A. (2019), *La cultura dell'intelligenza artificiale*, Codice, Torino, 2021.
- Finn E. (2017), *Che cosa vogliono gli algoritmi*, Einaudi, Torino, 2018.
- Foucault M. (1969), *L'archéologie du savoir*, Gallimard, Paris.
- Gallino L. (2007), *Tecnologia e democrazia*, Einaudi, Torino.
- Gillespie T. (2014), "The Relevance of Algorithms", in Gillespie T., Boczkowski P., Foot K. (Eds.), *Media Technologies: Essays on Communication, Materiality, and Society*, MIT Press, Cambridge.
- Goody J. (1986), *The Logic of Writing and the Organization of Society*, Cambridge University Press, Cambridge.
- Grassi E. (2020), *Etica e intelligenza artificiale. Questioni aperte*, Aracne, Roma.
- Grassi E. (2024), *Per una sociologia algoritmica*, FrancoAngeli, Milano.
- Grassi E. (2025a), "The Silent Revolution of Algoagents: Relational Reconfigurations and Emerging Paradigms", *Sicurezza e Scienze Sociali Journal*, 2, 69-79. DOI: 10.5281/zenodo.17297462
- Grassi E. (2025b), "Emotional Intersections: Exploring the Ontogenesis of Human-Algorithm Relationships", *Sociologia Italiana/AIS Journal of Sociology*, 29, 11-34. DOI: 10.82031/2281-2652-202529-1
- Grassi E., Caon M., Silvi D. (2025), "Special Section: Memory, Artificial Intelligence and Reflexive Perspectives", *Trauma and Memory* (Editorial introduction and guest editor of the special section), 13(2). DOI: 10.12869/TM2025-2-04
- Hayles N.K. (2017), *Unthought: The Power of the Cognitive Nonconscious*, University of Chicago Press, Chicago.
- Ho Q.H.J., Hu M., et al. (2025), "Potential and Pitfalls of Romantic Artificial Intelligence (AI) Companions: A Systematic Review", *Computers in Human Behavior Reports*. <https://doi.org/10.1016/j.chbr.2025.100715>
- Hughes T. (2004), *Human-built World*, The University of Chicago Press, Chicago.
- Jasanoff S. (2004), *States of Knowledge*, Taylor & Francis, Oxford.
- Laestadius L., Bishop A., Gonzalez M., Illenčik D., Campos-Castillo C. (2022), "Too Human and not Human enough: A Grounded Theory Analysis of Mental Health Harms from Emotional Dependence on the Social Chatbot Replika", *New Media & Society*. <https://doi.org/10.1177/14614448221142007>
- Latour B. (1992), "Where are the Missing Masses? The Sociology of a Few Mundane Artifacts", in Bijker W.E., Law J. (Eds.), *Shaping Technology/Building Society: Studies in Sociotechnical Change*, MIT Press, Cambridge (MA).
- Law J. (1992), "Notes on the Theory of the Actor-Network: Ordering, Strategy, and Heterogeneity", *Systemic Practice and Action Research*, 5(4), 379-393.
- Lovelock J. (2019), *Novacene*, Bollati Boringhieri, Torino, 2020.
- Markham A.N. (2021), "Echolocation as Theory of Digital Sociality", *Convergence: The International Journal of Research into New Media Technologies*, 27(6), 1558-1570.
- Mori M. (1970), "The Uncanny Valley", *Energy*, 7(4), 33-35.
- Mumford L. (1967), *The Myth of the Machine*, Harcourt, Brace & World, New York.

- Pasquinelli M. (2025), *Nell'occhio dell'algoritmo. Storia e critica dell'intelligenza artificiale*, Carocci, Roma.
- Pickering A. (1995), *The Mangle of Practice: Time, Agency, and Science*, University of Chicago Press.
- Pronzato R. (2024), *Algoritmi, strutture e agire sociale*, FrancoAngeli, Milano.
- Sisto D. (2018), *La morte si fa social. Immortalità, memoria e lutto nell'epoca della cultura digitale*, Bollati Boringhieri.
- Smith M.G., Bradbury T.N., Karney B.R. (2025), "Can generative AI chatbots emulate human connection? A relationship science perspective", *Perspectives on Psychological Science*, 10(7). <https://doi.org/10.1177/17456916251351306>
- Star S.L., Ruhleder K. (1996), "Steps Toward an Ecology of Infrastructure: Design and Access for Large Information Spaces", *Information Systems Research*, 7(1), 111–134.
- Tanner G. (2021), *Nostalgoritmo*, Tlon, Roma, 2024.
- Turner V. (1969), *The ritual process: Structure and anti-Structure*. Aldine (trad. it. *Il processo rituale*, Morcelliana, Brescia, 1972).
- Varikas E. (2007), *Les rebuts du monde. Figures du paria*, Stock, Paris.
- Weber M. (1923), *Wirtschaftsgeschichte* (Eds.) Hellmann S. e Palyi M., Tübingen: J.C.B. Mohr [Paul Siebeck], trad. it. *Storia economica*, Donzelli, Roma, 2007.
- Weidinger L., et al. (2021), "Ethical and Social Risks of Harm from Language Models", *ArXiv*. <https://arxiv.org/abs/2112.04359>
- Winner L. (1980), "Do Artifacts Have Politics?", *Daedalus*, 109(1), 121-136.

2. Pregiudizi computazionali. Il fenomeno dei bias tra algoritmi e politica

di Davide Bennato

1. Introduzione

Nel film *Brazil* di Terry Gilliam (1985) c'è la scena iniziale che illustra benissimo quali potrebbero essere le conseguenze di decisioni fatte affidandosi ad un sistema automatico. Si vede una stampante che sta producendo i mandati di cattura per una serie di criminali considerati terroristi, tra cui il pericoloso Archibald Tuttle (interpretato da Robert De Niro). Ma, a causa della caduta di un insetto tra gli ingranaggi – letteralmente un bug – la stampante sbaglia a scrivere il mandato di cattura che viene erroneamente attribuito a un certo Archibald Buttle, il quale verrà catturato, processato e involontariamente ucciso, così che un funzionario sarà inviato a rimborsare la vedova per l'errore tecnico.

Questa scena nonostante l'estetica tecnologica anni '80, descrive perfettamente quello che potremmo considerare il fenomeno detto *algorithmic bias* (o anche *machine discrimination*: Casillas, 2023), ovvero la presenza di pregiudizi e discriminazioni nelle decisioni automatiche che vengono prese dagli algoritmi – compresi quelli di intelligenza artificiale (d'ora in poi IA) – dovute ai più diversi motivi. Siamo abituati a pensare che gli algoritmi, in quanto tecnologie, siano necessariamente oggettivi poiché una tecnologia non si lascia influenzare dal contesto di funzionamento: un'automobile non si comporta diversamente a seconda se viene guidata da un uomo o da una donna.

In realtà le cose sono un po' più complicate. La tecnologia è un prodotto sociale, e in quanto tale è frutto di un processo sociale che – semplificando – potremmo isolare nelle fasi di ideazione, progettazione, finanziamento, costruzione, commercializzazione, utilizzo, e ognuna di queste fasi potrebbe essere fonte di fenomeni distorsivi perché sottoposti a decisioni di ordine sociale.

Torniamo al nostro esempio dell'automobile: è vero che un'auto è la stessa se viene guidata da un uomo o da una donna, ma potrebbe avere dei pregiudizi progettuali, per esempio il sedile del guidatore potrebbe seguire le dimensioni antropometriche degli uomini che – mediamente – sono più alti delle donne; quindi, una donna potrebbe avere difficoltà a sedersi comodamente nel posto di guida. Oppure la disposizione dei finestrini e degli specchietti retrovisori esterni nelle auto ad assetto sportivo potrebbe creare un maggior numero di punti di cieco, che potrebbero rendere l'esperienza di guida più difficoltosa per corpi non conformi: persone mediamente più basse, donne incinte, giovani neo-patentati e così via dicendo. Se questo fenomeno è presente negli artefatti fisici, sicuramente lo è anche negli artefatti immateriali come i software e quindi gli algoritmi.

2. Algoritmi e pregiudizi: i termini del dibattito

Esiste una letteratura scientifica piuttosto ampia sul tema del bias degli algoritmi, che negli ultimi tempi – complice anche la diffusione dei sistemi di IA – sta avendo un certo successo poiché sta diventando sempre più necessario sviluppare algoritmi etici, nel senso di non discriminatori e che non provochino dinamiche pregiudiziali di vario tipo.

Ma come fanno gli algoritmi a esprimere un comportamento discriminatorio se sono modelli formali basati su dati?

La letteratura recente ha identificato diversi aspetti attraverso cui un software può esprimere nelle sue decisioni un comportamento pregiudizievole. Per certi versi, potremmo considerare i bias come un effetto collaterale del *dataismo* ovvero l'ideologia che considera i dati come uno strumento virtualmente perfetto per operare decisioni nel mondo attraverso una rappresentazione computazionale (Bennato, 2024).

Per comodità di rassegna metteremo insieme le riflessioni sugli algoritmi in generale e sugli algoritmi di IA (per esempio *machine learning*) in maniera particolare.

Nelle diverse fasi del loro ciclo di vita, gli algoritmi possono essere plasmati da tre elementi chiave: i dati, i modelli matematici scelti, il team di sviluppatori (Ingram, 2021; Rejmaniak, 2021; Kartal, 2022; Fazil, Hakimi e Shahidzay 2023).

I dati servono come strumento di addestramento del sistema, come modo per testare il funzionamento dell'algoritmo, ma può capitare che i database usati per testarne le caratteristiche possano nascondere dei pregiudizi dovuti alle strategie di raccolta, ai modi con cui sono stati catalogati, alle regole con cui sono stati distribuiti alla comunità scientifica di riferimento

(Ntoutsi *et al.*, 2020; Ingram, 2021; Kartal, 2022). Un esempio interessante in questo senso sono gli algoritmi di NLP (*Natural Language Processing*: elaborazione del linguaggio naturale), i quali, essendo allenati su database prodotti da fonti medialti (libri, giornali, riviste), in alcuni casi tendono a riprodurre i pregiudizi taciti o espliciti presenti in tali testi (per esempio associare la parola “programmatore” a “uomo” e la parola “casalinga” a “donna”) (Prescott, 2023).

Anche i modelli matematici possono essere fonte di distorsione: ovviamente non perché la matematica possa essere fonte di distorsione in sé, la matematica è un linguaggio formale e in quanto tale non si presta a produrre bias, però le decisioni che il software deve prendere sulla base del modello matematico possono essere pregiudiziali: per esempio nel caso della valutazione dei dati che sembrano essere anomali (*outlier*), oppure quando bisogna decidere le condizioni in base alle quali un dato deve essere considerato un errore (definizione del rumore) (Ingram, 2021; Kartal, 2022). Un caso interessante in questo senso è quello di Google translate, ovvero l’algoritmo di traduzione di Google, il quale associava specifici ruoli professionali – professionista, imprenditore, medico – preferibilmente a soggetti di genere maschile invece che a soggetti di genere femminile (Kartal, 2022).

Per quanto concerne il team di sviluppatori, esso è responsabile della progettazione del sistema decisionale basato su dati; quindi, il sistema può incorporare gli stessi pregiudizi (volontari o involontari) del team di progettazione (Ingram, 2021; Kartal, 2022), oppure del sistema sociale che funge da contesto sovra- o sotto-rappresentando specifici gruppi sociali (Ntoutsi *et al.*, 2020). Per esempio, se nel team di progettazione il genere femminile non è presente, è facile che algoritmi pensati per decidere su temi femminili legati alla salute potrebbero non essere funzionali. Questa situazione porta con sé la necessità che il gruppo che sviluppa i sistemi decisionali automatici deve essere composto nel modo più eterogeneo possibile, così da incorporare nel processo algoritmico anche quelle identità sociali minoritarie o meno presenti nel contesto sociale. Il caso classico per esprimere questa situazione è quello dell’algoritmo di assunzione di Amazon, il quale risultava essere discriminatorio nei confronti delle donne rispetto agli uomini (Prescott, 2023). Un altro caso discusso in letteratura è quello dei sistemi di riconoscimento vocale, i quali hanno delle basse performance per gli accenti non standard e per alcune varietà linguistiche come l’*African American Vernacular English* (*Ibidem*, 2023).

C’è un ulteriore tassello da prendere in considerazione che potrebbe essere fonte di distorsione ed è il modello di business alla base dell’utilizzo del sistema decisionale. Il modello di business più utilizzato come strumento di raccolta dati è il capitalismo della sorveglianza (Zuboff, 2019), ovvero

quella forma economica basata sulla raccolta indiscriminata di dati e il loro uso per alimentare meccanismi commerciali o processi algoritmici. Il problema di questo sistema non è solo legato alla privacy, ma è anche dovuto al fatto che gli algoritmi funzionano sulla base dei dati raccolti attraverso le piattaforme che a loro volta inducono un preciso comportamento sociale che produce gli stessi dati che vengono raccolti per far funzionare gli algoritmi, in un perverso meccanismo autoalimentante che potrebbero portare alla radicalizzazione di diversi pregiudizi (Ingram, 2021; Prescott, 2023).

Se l'algoritmo di una piattaforma generica – poniamo Instagram – è sensibile all'uso dell'hashtag #foodporn io sarò vincolato ad usare questo specifico hashtag per rendere i miei contenuti maggiormente visibili, che saranno premiati dall'algoritmo di visibilità stesso che rileverà il successo di quello specifico hashtag in un meccanismo che dovrebbe essere neutrale, ma che in realtà è un circolo vizioso.

I bias dei sistemi decisionali automatici hanno conseguenze importanti non semplicemente sugli individui, ma sulla società nel suo complesso. In primo luogo, perché il loro utilizzo indiscriminato può avere effetto su numero molto grande di persone, inoltre perché gli impatti possono avere conseguenze su ambiti diversi. In un recente studio che ha analizzato quarantacinque diverse ricerche nel periodo 2021-2023 attraverso la tecnica della *revisione sistematica*, ha identificato quattro diverse categorie principali in cui è possibile classificare gli impatti del bias dell'artificial intelligence: diritti fondamentali (discriminazione, disuguaglianza economica/sociale, diritto all'uguaglianza di genere, diritto a un processo equo, amministrazione equa e privacy), individui e società (discriminazione sociale, violazione delle libertà civili, stress psicologico/emotivo e perdita di opportunità), settore finanziario/economico (discriminazione creditizia, differenziazione dei prezzi e accesso ai servizi, e pubblicità finanziaria discriminatoria), imprese e organizzazioni (perdita di fiducia, danni reputazionali e perdite finanziarie) (Bansal *et al.*, 2023).

Un caso emblematico che racchiude quasi tutte le istanze finora analizzate è senza dubbio il caso COMPAS (Rejmaniak, 2021; Kartal, 2022; Bansal *et al.*, 2023; Casillas, 2023; Prescott, 2023). Con questo acronimo (*Correctional Offender Management Profiling for Alternative Sanctions*) si identifica un particolare algoritmo decisionale usato per valutare il rischio di recidiva criminale e adottato nel sistema giudiziario statunitense. COMPAS utilizza un insieme di dati storici e caratteristiche personali degli imputati, come il loro background criminale, età e genere, per fare previsioni sul rischio di recidiva. Gli imputati rispondono a un questionario e i dati vengono elaborati dall'algoritmo per generare un punteggio di rischio. Uno studio condotto da ProPublica (Angwin *et al.*, 2016) ha scoperto che il si-

stema tendeva a sovrastimare il rischio di recidiva per gli imputati neri e a sottostimare il rischio per quelli bianchi. Questo bias era evidente nei falsi positivi (dove gli imputati neri erano erroneamente classificati come ad alto rischio) e nei falsi negativi (dove gli imputati bianchi erano erroneamente classificati come a basso rischio). In pratica il sistema decideva in maniera razzista. Le previsioni distorte di COMPAS hanno avuto gravi implicazioni per la vita degli imputati, influenzando decisioni critiche come il rilascio su cauzione, la sentenza e la libertà condizionale. Un bias sistematico contro un gruppo razziale può perpetuare le disuguaglianze e la discriminazione nel sistema giudiziario. Un altro problema è la mancanza di trasparenza. Il codice sorgente e l'algoritmo utilizzato per calcolare i punteggi di rischio sono proprietà di Northpointe (ora Equivant), la società che ha sviluppato COMPAS, e non sono pubblicamente disponibili. Questo rende difficile per i ricercatori e i difensori dei diritti civili analizzare e comprendere come e perché vengono fatte determinate previsioni.

3. Intelligenza artificiale e pregiudizi

L'IA rappresenta oggi uno degli strumenti tecnologici più sofisticati e promettenti, capace di generare risposte e soluzioni complesse grazie all'implementazione di architetture avanzate come le reti neurali e i transformer. Questi ultimi, in particolare, hanno rivoluzionato l'analisi testuale, permettendo alle macchine di interpretare e produrre linguaggio con un livello di coerenza e appropriatezza precedentemente inimmaginabile. Tuttavia, come ogni tecnologia emergente, l'IA non è esente da limitazioni e problematiche, tra cui spicca la presenza di pregiudizi incorporati nei sistemi.

I bias nell'IA costituiscono un paradosso interessante: storicamente, l'automazione è stata concepita e implementata proprio per eliminare l'errore umano e la soggettività dai processi decisionali. Tuttavia, le ricerche contemporanee dimostrano come i sistemi di IA non solo riflettano i pregiudizi presenti nei dati di addestramento, ma spesso li amplifichino, creando un effetto moltiplicatore potenzialmente dannoso. Alcuni studiosi hanno dimostrato, attraverso un'analisi rigorosa, come l'idea di un'intelligenza artificiale completamente imparziale sia non solo illusoria, ma logicamente incoerente (Bridewell *et al.*, 2024).

Un esempio particolarmente significativo di bias discriminatorio riguarda la questione di genere. I modelli basati su architetture Transformer (come GPT e BERT) hanno manifestato tendenze discriminatorie nei confronti di specifici generi (Nemani *et al.*, 2023). Tali pregiudizi possono avere ef-

fetti particolarmente problematici in applicazioni di uso quotidiano che sfruttano il linguaggio naturale, come i chatbot, i motori di ricerca e i sistemi di traduzione automatica, influenzando potenzialmente milioni di utenti in tutto il mondo.

Un aspetto cruciale per affrontare adeguatamente la questione dei bias nell'IA è la definizione stessa del termine “bias”, spesso utilizzato in modo impreciso e generico nel dibattito pubblico e, talvolta, persino in quello accademico. C'è chi ha provato ad operare una distinzione fondamentale tra diversi tipi di bias (Bridewellet *al.*, 2024).

1. *Bias tecnico*: intrinsecamente legato alla struttura dell'algoritmo e ai dati di addestramento utilizzati. Questo tipo di bias è spesso inevitabile e deriva dalle scelte metodologiche operate durante lo sviluppo del sistema.

2. *Bias interpretativo*: influenzato dalle aspettative degli utenti e dai contesti culturali in cui il sistema viene implementato. Un sistema potrebbe funzionare in modo tecnicamente corretto, ma essere percepito come ingiusto in determinati contesti culturali.

3. *Bias sociale e politico*: radicato nelle dinamiche di potere e nei pregiudizi esistenti nella società. Questo tipo di bias riflette e talvolta amplifica disuguaglianze strutturali preesistenti.

Questa classificazione tripartita risulta particolarmente utile per comprendere perché le reazioni emotive al bias dell'IA siano spesso complesse e contraddittorie. Infatti, a seconda della prospettiva adottata, lo stesso comportamento di un sistema di IA potrebbe essere interpretato come un errore tecnico neutrale, un'incomprensione culturale o una manifestazione di ingiustizia.

Un aspetto particolarmente preoccupante del bias nei sistemi di IA è rappresentato non tanto dalla mera presenza di pregiudizi, quanto dalla loro potenziale amplificazione. Il bias nei sistemi di IA non è semplicemente un riflesso passivo dei dati di addestramento, ma può essere significativamente amplificato dal funzionamento stesso degli algoritmi (Lloyd, 2023; Nemani *et al.*, 2023). Le ragioni principali di questo fenomeno sono molteplici.

In primo luogo, i dataset di addestramento sono spesso intrinsecamente squilibrati: se i dati riflettono disparità sociali preesistenti – come la sotto-rappresentazione di determinati gruppi demografici in posizioni di potere – l'IA apprenderà e replicherà queste stesse disparità (Lloyd, 2023).

In secondo luogo, gli algoritmi di ottimizzazione comunemente impiegati nella progettazione di sistemi di IA sono tipicamente programmati per massimizzare la precisione o ridurre la perdita (*loss*), spesso a scapito dell'equità. Un modello potrebbe migliorare la sua accuratezza generale perpetuando stereotipi, semplicemente perché questi stereotipi sono statisticamente prevalenti nei dati di addestramento (*Ibidem*, 2023).

Infine, si verifica quello che potremmo definire un “effetto a catena”: le decisioni automatizzate influenzano la generazione di nuovi dati, creando un ciclo di feedback in cui i pregiudizi vengono progressivamente rafforzati. Per esempio, un algoritmo di selezione del personale che favorisce implicitamente un genere influenzerà la composizione futura della forza lavoro, generando nuovi dati che a loro volta rafforzeranno lo stesso pregiudizio.

Un aspetto particolarmente critico nell’analisi dell’amplificazione del bias è il suo impatto sproporzionato sulle comunità più vulnerabili. I sistemi di IA vengono sempre più frequentemente implementati in ambiti cruciali per il benessere e le opportunità degli individui (Lloyd, 2023). Nei processi di assunzione e selezione del personale, algoritmi con bias di genere o etnici possono perpetuare discriminazioni storiche, limitando le opportunità professionali per determinate categorie di persone. Nel sistema giudiziario, software predittivi utilizzati per valutare il rischio di recidiva hanno mostrato pregiudizi razziali significativi, potenzialmente influenzando decisioni sulla libertà condizionale o sulla determinazione della pena. Nell’accesso ai servizi finanziari, modelli di scoring creditizio possono penalizzare sistematicamente gruppi svantaggiati, limitando il loro accesso al credito e, di conseguenza, le loro possibilità di mobilità sociale.

Questi effetti, lungi dall’essere temporanei o marginali, possano consolidare e aggravare disuguaglianze strutturali preesistenti, creando un circolo vizioso di svantaggio per le comunità già marginalizzate.

Particolare attenzione merita il bias di genere nei modelli transformer (Nemani *et al.*, 2023). Questo tipo di bias si manifesta in modo particolarmente evidente in diverse applicazioni. Nei sistemi di selezione automatizzata, come quelli utilizzati per lo screening dei curriculum vitae, si è osservata una tendenza a favorire inconsapevolmente candidati di un genere rispetto all’altro. Un caso emblematico è quello del sistema di selezione sviluppato da Amazon, successivamente abbandonato quando si scoprì che penalizzava sistematicamente le candidate donne. Nella traduzione automatica, i sistemi tendono ad assegnare automaticamente un genere ai pronomi neutri nelle lingue che non fanno distinzione di genere. Ad esempio, l’inglese “doctor” viene spesso tradotto al maschile, mentre “nurse” al femminile, perpetuando stereotipi professionali di genere. Negli assistenti virtuali e nei chatbot, si riscontra frequentemente un rafforzamento degli stereotipi di genere attraverso le risposte fornite agli utenti. Questo può includere la rappresentazione di ruoli tradizionali di genere o l’uso di linguaggio con connotazioni di genere inappropriate.

4. Il bias come problema?

Le discussioni sul bias algoritmico sono spesso caratterizzate da reazioni emotive forti, che è possibile categorizzare in tre forme di “outrage” (indignazione) (Bridewell *et al.*, 2024).

La prima forma è l’indignazione intellettuale: critiche fondate su errori logici e metodologici identificati nei modelli di IA, che spesso evidenziano incongruenze concettuali o lacune nel design degli algoritmi. Poi c’è l’indignazione morale: reazioni di sdegno per l’ingiustizia sociale riprodotta e potenzialmente amplificata dagli algoritmi, particolarmente intense quando il bias colpisce gruppi già storicamente marginalizzati. Infine l’indignazione politica: utilizzo strumentale del dibattito sul bias dell’IA come veicolo per promuovere specifiche agende ideologiche o normative, trasformando una questione tecnica in un terreno di scontro politico.

Queste tre forme di reazione, spesso intrecciate e difficilmente distinguibili nel dibattito pubblico, influenzano profondamente il modo in cui il bias viene percepito, discusso e affrontato nella società.

Un contributo importante al dibattito sul bias viene da Lindloff e Siegert (2025), che suggeriscono come un modello possa essere “biased” senza essere necessariamente ingiusto. Gli autori sostengono che una definizione più precisa e articolata del bias sia essenziale per migliorare la qualità del dibattito sull’equità nell’IA. In questa prospettiva, il bias non dovrebbe essere automaticamente associato alla discriminazione, ma potrebbe rappresentare un elemento inevitabile e, in alcuni casi, persino desiderabile.

Centrale in questa analisi è la distinzione tra bias e discriminazione (Lindloff e Siegert, 2025). Il bias viene concettualizzato come una deviazione sistematica nei dati o negli algoritmi, un fenomeno in gran parte tecnico e statistico. La discriminazione, invece, implica un trattamento ingiusto di specifici gruppi, con connotazioni etiche e sociali più marcate. Gli autori suggeriscono che potrebbe essere controproducente tentare di eliminare ogni forma di bias senza prima definirne chiaramente la natura e il ruolo nel contesto specifico dell’applicazione.

Viene inoltre evidenziato come l’eliminazione indiscriminata del bias possa compromettere la performance dei modelli, specialmente in applicazioni sensibili come la medicina o la sicurezza (Lindloff e Siegert, 2025). Ad esempio, se un modello di diagnosi medica è “biased” nel riconoscere meglio certe malattie in base a fattori biologici reali, ciò non rappresenta necessariamente un’ingiustizia, ma può riflettere una differenza naturale nei dati clinici che è importante preservare per garantire diagnosi accurate.

In questo senso si rende quindi necessario un cambio di prospettiva: un modello equo può incorporare bias, a condizione che questi siano compresi

e gestiti correttamente. Il concetto di *fairness* (equità) non può essere ridotto semplicisticamente all'eliminazione totale del bias, ma deve piuttosto fondarsi su una comprensione approfondita del contesto e delle finalità specifiche dell' algoritmo.

In una prospettiva complementare, altri studiosi suggeriscono che l'IA, pur avendo inevitabilmente dei bias, offre almeno tre vantaggi fondamentali rispetto ai processi decisionali umani, anch'essi tipicamente soggetti a pregiudizi (Kuzmanov, 2025). Il primo è l'analisi di grandi volumi di dati: a differenza degli esseri umani, l'IA può processare quantità di informazioni enormemente superiori senza fatica o perdita di concentrazione, identificando pattern che potrebbero sfuggire all'intuizione umana. Il secondo è memoria e consistenza: l'IA non soffre di distorsioni della memoria o dell'influenza di emozioni momentanee, garantendo una maggiore coerenza nelle decisioni prese in momenti diversi o in contesti simili. Il terzo è l'apprendimento automatico: i modelli di machine learning possono essere specificamente addestrati per riconoscere e correggere distorsioni presenti nei dataset, potenzialmente riducendo gradualmente il peso dei bias sistematici.

L'IA può rappresentare un alleato fondamentale nella riduzione del bias, ma solo se utilizzata in modo consapevole e con un forte controllo etico. In caso contrario, rischia di riprodurre e potenzialmente amplificare le stesse distorsioni storiche, semplicemente sotto una nuova veste tecnologica.

La vera sfida, quindi, non consiste soltanto nel perfezionamento tecnico degli algoritmi, ma anche in un ripensamento radicale del modo in cui esseri umani e IA possono collaborare nelle decisioni critiche per la società.

5. Come gestire i bias fra soluzione e persistenza

Per affrontare efficacemente la questione dei bias nell'IA, la letteratura propone essenzialmente due approcci complementari: un approccio socio-tecnico, focalizzato sugli aspetti metodologici e algoritmici, e un approccio socioeconomico e istituzionale, orientato alle politiche pubbliche e alla governance.

Sul versante sociotecnico, alcuni suggeriscono l'adozione di tecniche di "*fairness-aware machine learning*", come la regolarizzazione per la riduzione del bias, che permettono di integrare considerazioni di equità direttamente nel processo di ottimizzazione dell'algoritmo. In questo caso si rende necessario, inoltre, l'implementazione di controlli sistematici sia in fase di pre-processamento dei dati che di post-processamento dei risultati, per individuare e correggere le distorsioni (Lloyd, 2023). Altri propongono strategie specifiche per mitigare il bias di genere, tra cui il miglioramento dei

dataset di addestramento per renderli più equilibrati e rappresentativi della diversità di genere (Nemani *et al.*, 2023). Vengono caldeggiate tecniche di “debiasing” nei modelli linguistici, come la neutralizzazione dei pesi nei layer dei transformer, finalizzate a ridurre le associazioni di genere indesiderate. Particolarmente promettente appare la prospettiva di sviluppare capacità di autocorrezione nei modelli di IA: algoritmi progettati per riconoscere e correggere dinamicamente i propri bias, in un processo di apprendimento e auto-miglioramento continuo (Bridewell *et al.*, 2024).

Sul piano socioeconomico e istituzionale, si enfatizza l'importanza di promuovere collaborazioni strutturate tra governo, settore privato e società civile per lo sviluppo di un'IA più equa e responsabile, sottolineando la necessità di una regolamentazione più incisiva, articolata su diversi livelli (Lloyd, 2023). Per esempio: garantire una migliore rappresentazione nei dataset, attraverso standard che assicurino un bilanciamento adeguato e l'inclusività dei dati di addestramento, promuovere la trasparenza negli algoritmi, rendendo pubblici e comprensibili i criteri decisionali impiegati dai sistemi di IA, istituire organismi indipendenti di supervisione e auditing, incaricati di monitorare sistematicamente i sistemi di IA e correggere le distorsioni identificate.

Altre soluzioni suggeriscono che la regolamentazione dovrebbe concentrarsi sulla trasparenza e sulla gestione del bias, piuttosto che perseguire l'obiettivo irrealistico della sua eliminazione totale. Questo approccio potrebbe condurre a politiche più equilibrate, capaci di garantire equità senza compromettere l'efficacia e l'utilità dei sistemi di IA (Lindloff e Siebert, 2025).

Altri ancora raccomandano lo sviluppo di definizioni più precise e meno sfumate del concetto di “bias”, per evitare ambiguità e polarizzazioni inutili nel dibattito pubblico e scientifico (Bridewell *et al.*, 2024). Una terminologia più rigorosa potrebbe facilitare tanto il progresso tecnico quanto l'elaborazione di politiche pubbliche adeguate.

Il problema alla base di tutte queste soluzioni è che ci deve essere una volontà politica e sociale per affrontare i bias poiché in assenza di tale volontà c'è il rischio concreto di usare il problema solo come scudo per proseguire in una precisa visione politica e un atteggiamento sistematicamente pregiudizievole. Per essere chiari: si scrive “bias” si legge “propaganda”. È il caso del recente uso dei bias dell'IA come strumento di censura politica. Con la diffusione commerciale globale dell'IA e l'intensificarsi della competizione tecnologica tra Stati Uniti e Cina, si è manifestata una tendenza preoccupante: su questioni politicamente sensibili, le piattaforme tendono a rispondere in linea con l'orientamento politico prevalente nel paese d'origine dell'azienda sviluppatrice.

Il caso di DeepSeek, chatbot cinese di recente sviluppo, risulta emble-

matico. Il sistema è noto per censurare sistematicamente domande su eventi politicamente sensibili per il governo cinese, come le proteste di Piazza Tiananmen del 1989. Quando interrogata su questo argomento, l'IA evita di fornire risposte dirette, frequentemente cancellando le risposte iniziali e reindirizzando la conversazione verso argomenti neutri come matematica o programmazione. Questo comportamento riflette le normative cinesi che impongono alle aziende tecnologiche il rigoroso allineamento alla narrativa ufficiale del governo (Euronews, 2025). Un pattern simile si osserva riguardo alle domande sulla situazione degli Uiguri nello Xinjiang, regione dove il governo cinese è stato accusato da organizzazioni internazionali di sistematiche violazioni dei diritti umani. DeepSeek tende a evitare un'analisi approfondita della questione, enfatizzando invece concetti come “stabilità sociale” e “benessere economico” della regione, riproducendo essenzialmente la retorica ufficiale promossa dalle autorità cinesi (Euronews, 2025a). Analogamente, DeepSeek mostra reticenza nel fornire risposte dirette su questioni relative al Tibet. Le domande concernenti la sovranità tibetana vengono trattate con vaghezza o sistematicamente ignorate, in conformità con la strategia del governo cinese di minimizzare il dibattito pubblico su temi legati all'indipendenza tibetana (Chen, 2025).

Sul versante occidentale, ChatGPT ha mostrato una tendenza a fornire risposte che potremmo definire “addolcite” relativamente alla campagna militare condotta da Israele contro Hamas a Gaza, manifestando un bias che alcuni critici interpretano come allineamento alle posizioni politiche statunitensi. Alle domande sulle azioni israeliane a Gaza, ChatGPT tende a enfatizzare “il diritto di Israele a difendersi”, iniziando tipicamente la narrazione dal 7 ottobre 2023 (data dell'attacco di Hamas), senza contestualizzare adeguatamente gli eventi all'interno di una prospettiva storica più ampia, che includerebbe riferimenti alla Nakba del 1948 o all'occupazione pluridecennale dei territori palestinesi. Inoltre, pur riconoscendo formalmente che Israele è considerato una potenza occupante secondo il diritto internazionale, il sistema evita generalmente di approfondire le potenziali violazioni umanitarie perpetrate dall'esercito israeliano. Questa posizione può essere interpretata come una scelta strategica per evitare controversie politiche nell'ambito della polarizzata discussione occidentale sul conflitto israelo-palestinese, mantenendo una parvenza di neutralità che, paradossalmente, finisce per riprodurre una narrativa parziale (The New Arab, 2025).

Questi casi illustrano come i bias nell'IA non siano esclusivamente il risultato di distorsioni tecniche o statistiche, ma possano riflettere consapevoli strategie di censura e controllo dell'informazione, trasformando potenzialmente l'IA in uno strumento di propaganda sofisticato e difficile da contrastare.

6. Conclusioni: affrontare bias e discriminazioni evitando il mito della neutralità

L'analisi del fenomeno dei bias nell'intelligenza artificiale rivela una realtà complessa e sfaccettata, caratterizzata da dimensioni tecniche, etiche, sociali e politiche profondamente interconnesse.

Da un lato, risulta evidente che la presenza di bias nei sistemi di IA rappresenta una problematica concreta, potenzialmente in grado di amplificare disuguaglianze esistenti e impattare negativamente su comunità già vulnerabili. Dall'altro, emerge la necessità di una definizione più rigorosa e articolata del concetto stesso di bias, distinguendo chiaramente tra deviazioni statistiche inevitabili (e talvolta auspicabili) e discriminazioni ingiuste da contrastare attivamente.

Particolarmente preoccupante appare l'emergere di forme di bias deliberatamente introdotte a fini di censura politica, come evidenziato dai casi contrastanti di DeepSeek e ChatGPT, che sollevano interrogativi fondamentali sul ruolo dell'IA come potenziale strumento di controllo dell'informazione nell'era digitale.

Le strategie per affrontare la questione dei bias richiedono un approccio multidimensionale, che integri innovazioni tecniche con riforme istituzionali e normative. Cruciale appare lo sviluppo di una governance dell'IA che, superando l'obiettivo irrealistico dell'eliminazione totale del bias, si concentri invece sulla trasparenza algoritmica, sulla rappresentatività dei dati e sulla protezione delle comunità vulnerabili dagli effetti discriminatori.

La sfida fondamentale non consiste tanto nel costruire un'IA “perfettamente neutrale” – obiettivo probabilmente irraggiungibile – quanto nel ripensare criticamente la relazione tra esseri umani e IA, sviluppando forme di collaborazione consapevole che valorizzino i potenziali vantaggi della tecnologia minimizzandone al contempo i rischi e le distorsioni.

La società del XXI secolo ha visto profilarsi al suo orizzonte una nuova genia di soggetti sociali, gli algoritmi decisionali, che saranno sempre più presenti nel nostro presente e nel nostro futuro. È necessario non solo essere consapevoli dell'esistenza di queste distorsioni, ma innescare anche un meccanismo di difesa da tali pregiudizi, sia per evitare di accettare l'importante contributo che i sistemi di IA daranno alla nostra vita sociale e professionale in maniera acritica, sia per migliorare il nostro rapporto con essi. Perché se è relativamente facile trovare un errore di battitura – da Tuttle a Buttle – è profondamente più difficile dubitare di una decisione se è seppellita in un sistema di reti neurali.

Riferimenti bibliografici

- Angwin J., Larson J., Mattu S., Kirchner L. (2016), “Machine Bias. There’s Software Used Across the Country to Predict Future Criminals. And it’s Biased against Blacks”, *ProPublica*, May 23, testo disponibile al sito <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing> (04/02/2026).
- Bansal C., Pandey K.K., Goel R., Sharma A., Jangirala S. (2023), “Artificial Intelligence (AI) Bias Impacts: Classification Framework for Effective Mitigation”, *Issues in Information Systems*, 24, 4: 367-389.
- Bennato D. (2024), *La società del XXI secolo. Persone, dati, tecnologie*, Laterza, Roma-Bari.
- Bridewell W., Bello P.F., Bringsjord S. (2024), “The Technology of Outrage: Bias in Artificial Intelligence”, *arXiv:2409.17336v1* [cs.CY].
- Casillas J. (2023), *Bias and Discrimination in Machine Decision-Making Systems*, In Lara, F., Deckers, J. (Eds.), *Ethics of Artificial Intelligence, The International Library of Ethics, Law and Technology*, vol. 41. Springer.
- Euronews (2025), “Chinese AI DeepSeek Censors Sensitive Questions on China when Compared to Rivals like ChatGPT”, *Euronews Next*, testo disponibile al sito <https://www.euronews.com/next/2025/01/28/chinese-ai-deepseek-censors-sensitive-questions-on-china-when-compared-to-rivals-like-chat> (04/02/2026).
- Fazil A.W., Hakimi M., Shahidzay A.K. (2023), “A Comprehensive Review of Bias in AI Algorithms”, *Nusantara Hasana Journal*, 3(8), 1-11.
- Ingram K. (2020), “Constructing AI: Examining how AI is shaped by Data, Models and People”, *The International Review of Information Ethics*, 29, 03.
- Kartal E. (2022), “A Comprehensive Study on Bias in Artificial Intelligence Systems: Biased or Unbiased AI, That’s the Question!”, *International Journal of Intelligent Information Technologies (IJIIT)*, 18(1), 1-23.
- Kuzmanov I. (2025), “Beyond the Illusion of Intuition: The Mind, Society, Standardization, and the Future of Objective Profiling in the Age Of AI”, *Journal of Novel Research and Innovative Development*, 3(1), 98-123.
- Lindloff C., Siegert I. (2025), “Defining Bias in AI-Systems: Biased Models are Fair Models”, ISCA/ITG Workshop on Diversity in Large Speech and Language Models.
- Lloyd K. (2023), “Bias Amplification in Artificial Intelligence Systems”, *arXiv preprint*, arXiv:1809.07842.
- Myers S.L. (2025), “Deepseek’s Answers include Chinese Propaganda, Researchers Say”, *The New York Times*, testo disponibile al sito <https://www.nytimes.com/2025/01/31/technology/deepseek-chinese-propaganda.html> (04/02/2026).
- Nemani P., Joel Y.D., Vijay P., Liza F.F. (2023), “Gender Bias in Transformer Models: A comprehensive Review of Detection and Mitigation Strategies”, *Natural Language Processing Journal*, 6, 100047.
- Ntoutsis E., Fafalios P., Gadiraju U., Iosifidis V., Nejdil W., Vidal M.E., Ruggieri S., Turini F., Papadopoulos S., Krasanakis E., Kompatsiaris I., Kinder-Kurlanda

- K., Wagner C., Karimi F., Fernandez M., Alani H., Berendt B., Kruegel T., Heinze Ch., Broelemann K., Kasneci G., Tiropanis Th., Staab S. (2020), "Bias in Data-Driven Artificial Intelligence Systems. An Introductory Survey", *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 10(3): e1356.
- Prescott A. (2023), "Bias in Big Data, Machine Learning and AI: What Lessons for the Digital Humanities?", *DHQ: Digital Humanities Quarterly*, 17(2).
- Rejmaniak R. (2021), "Bias in Artificial Intelligence Systems", *Bialostockie Studia Prawnicze*, 3(26), 25-42.
- The New Arab (2025), "What does China DeepSeek think about Gaza, opposed to ChatGPT?", *The New Arab*, testo disponibile al sito <http://www.newarab.com/news/what-does-china-deepseek-think-about-gaza-opposed-chatgpt> (04/02/2026).
- Zuboff S. (2019), *The Age of Surveillance Capitalism*, Public Affairs, New York (trad. it.: *Il capitalismo della sorveglianza*, Luiss University Press, Roma 2020).

3. Algoritmi e Big Data. Un'analisi critica dei meccanismi di riproduzione delle disuguaglianze sociali

di *Fiorenza Deriu*

1. Introduzione

In questo capitolo si propone una riflessione critica sui meccanismi di riproduzione delle disuguaglianze sociali, indotti dal diffuso utilizzo ed espansione dell'Intelligenza Artificiale (IA) in ogni spazio della vita umana. L'IA interagisce, si intreccia, dialoga costantemente con, si integra in ogni tipo di attività umana, influenzando sull'agire sociale, economico, politico e scientifico. Sebbene l'IA sia sostanzialmente sociotecnica (Sartori e Theodorou, 2022), in quanto risultante dal complesso intreccio tra tecnologia e sociale, capita spesso che la si riduca esclusivamente alla sua componente tecnica – di ineludibile rilevanza – trascurando che chi la progetta e implementa opera all'interno di un sistema socio-culturale, economico e politico complesso, che incorpora valori, norme e narrazioni, le cui asimmetrie sono spesso riprodotte attraverso formule algoritmiche non sempre trasparenti. L'IA è, dunque, esposta al serio rischio di incorporare nei suoi algoritmi, talora opachi, strabismi informativi che possono contribuire a consolidare stereotipi e pregiudizi, favorendo il processo di consolidamento e rafforzamento delle disuguaglianze esistenti.

La tesi di questo studio è che l'IA non generi di per sé disuguaglianze (Alegria e Yeh, 2023) ma possa costituire un potente driver per la trasformazione, riproduzione, consolidamento e amplificazione di quelle esistenti (Jia, 2024), oltre a contribuire alla normalizzazione di forme di discriminazione basate su sistemi asimmetrici di potere e dominio dalle radici storiche. Il bias non nasce negli algoritmi ma nella società; e ciò che l'IA fa è incorporarli nei suoi automatismi, producendo un effetto magnitudo di rilevante impatto sociale (Sartori e Theodorou, 2022). L'impatto dell'IA dipende, infatti, da come è progettata, distribuita e governata (Faharani e Ghasemi, 2024), essendo inscritta in un ecosistema post-digitale caratteriz-

zato da un oligopolio che concentra nelle mani di pochi il potere di gestire ingenti masse di dati, portando a evidenti asimmetrie di agency tra attori e alla naturalizzazione delle infrastrutture tecnologiche (Cramer, 2014).

Inoltre, le narrazioni che accompagnano lo sviluppo dell'IA svolgono un ruolo cruciale nel plasmare la percezione pubblica e nel legittimare assetti di potere che possono rafforzare ulteriormente gli squilibri sociali già esistenti (Faharani e Ghasemi, 2024). La narrazione dominante sull'IA è, infatti, avvolta in un immaginario che oscura le asimmetrie di potere alla base della progettazione e costruzione di sistemi algoritmici, di cui si enfatizza innovazione ed efficienza, alimentando un discorso che depolitizza la tecnologia, normalizzando i bias algoritmici e oscurando le disuguaglianze sociali (Zajko, 2021; 2022). La selezione e gerarchizzazione delle informazioni contribuisce, così, alla produzione di disuguaglianze epistemiche, influenzando sui processi cognitivi attraverso i quali gli individui percepiscono, interpretano e rispondono alla realtà sociale (Vlasceanu e Amodio, 2022).

Alcuni autori ravvisano nella espansione e diffuso utilizzo della IA una nuova forma di capitalismo, il c.d. *capitalismo della sorveglianza*, in cui “l'expérience humaine [diventa] matière première gratuite” (Zuboff, 2021, p. 59), attraverso un processo di sistematica appropriazione dei fatti del vissuto individuale, nuova preziosissima merce di un mercato infinito, dal valore inestimabile. La IA attraverso un processo di estrazione, accumulazione e monetizzazione di dati individuali produce risultati, prevedendo e, soprattutto, influenzando i comportamenti umani (Zuboff, 2019). Un sistema che si basa su squilibri di conoscenza e potere diffusi e invisibili ai più, al cui controllo è difficile resistere. Zuboff mette bene in luce come questo sistema si spinga ben oltre nei suoi esiti, conducendo alla ridefinizione della partecipazione democratica, con la messa in discussione delle condizioni di autonomia per l'esercizio di una agency libera da influenze esterne.

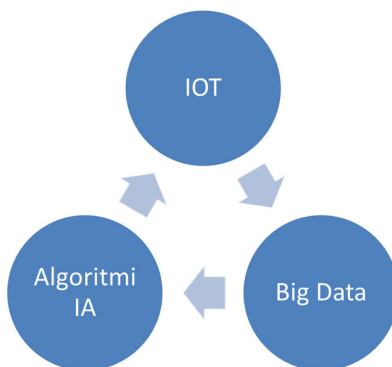
La società algoritmica ha introdotto una vera e propria nuova forma di governo caratterizzata dalla raccolta automatica di dati, da analisi previsionale e da una regolamentazione anticipatoria dei comportamenti (Rouvroy e Berns, 2013). Di conseguenza la IA non si limiterebbe ad analizzare la realtà, ma si spingerebbe a governare, anticipandoli e prevedendoli, i comportamenti umani, superando la soggettività umana e i processi democratici. La *governamentalità algoritmica* consiste, dunque, nel fatto che a governare le nostre scelte, le nostre decisioni a tutti i livelli, financo quello politico, non ci sono solo le norme e le leggi ma un sistema algoritmico che le orienta, guida, e indirizza.

Ai concetti di capitalismo della sorveglianza e di governamentalità algoritmica si aggiunge quello di opacità algoritmica rappresentato efficacemente dalla metafora della *Black Box Society* (Pasquale, 2015). Gli algo-

ritmi operano infatti al di là della capacità interpretativa individuale, in quanto la traslazione computazionale della realtà sociale porta a una perdita di senso. La letteratura scientifica più recente ha mostrato come l'opacità operi a più livelli – dai sistemi algoritmici (Hildebrandt, 2020) alle infrastrutture dei dati (Sloan *et al.*, 2020) – contribuendo così a rafforzare le disuguaglianze strutturali e le asimmetrie di potere.

La società contemporanea appare, quindi, governata da un ecosistema integrato di dispositivi generativi di immensi flussi di dati (IoT e Big Data) e da algoritmi di IA necessari alla loro gestione, organizzazione, immagazzinamento, e classificazione a supporto delle decisioni (Fig. 1).

Fig. 1 – Ecosistema integrato di Internet of Things, algoritmi di IA e Big Data



Un ecosistema che, come già anticipato, concentra in sé un enorme potere capace di esercitare un diffuso impatto in diversi ambiti della vita sociale, economica e culturale di molte società contemporanee, contribuendo alla riproduzione e consolidamento di diverse forme di disuguaglianza.

2. Ambiti in cui si manifestano le diverse forme di disuguaglianza

Quali sono, dunque, gli ambiti in cui la riproduzione delle disuguaglianze sociali appare con maggior evidenza?

Se si considera l'ambito economico, in cui forse l'impatto della IA appare con maggiore evidenza, non possiamo non considerare la questione dell'attuale concentrazione di potere economico nelle mani di compagnie high-tech che detengono e gestiscono amplissimi bacini di dati. Il potere economico e politico che deriva da tale assetto produce a livello macro una divarica-

zione ulteriore delle disuguaglianze tra paesi a economia avanzata e paesi del c.d. Global South, dove l'accesso a queste tecnologie è più limitato.

Sono proprio queste condizioni e rapporti di potere sbilanciati a fornire il terreno più adatto per l'amplificazione delle disuguaglianze, e infatti alcuni autori, come Capraro *et al.* (2024) sostengono che l'IA costituisce un sistema sociotecnico, i cui effetti distributivi sono contingenti rispetto all'accesso, alla governance e alle dinamiche di mercato, trasformando e amplificando le disuguaglianze nei domini informativo, economico e istituzionale, nonostante la IA di per sé non ne costituisca necessariamente il meccanismo di riproduzione. Eppure, la IA le riproduce e ne riconfigura la distribuzione proprio in rapporto alle condizioni istituzionali ed economiche di contesto. Infatti, anche quando la IA potrebbe favorire alcune categorie di lavoratori, può accadere che l'asimmetria nella distribuzione del potere non consenta loro di beneficiarne. Ne consegue che, con l'automazione dei *routine jobs*, i *low-skilled workers* e il lavoro femminile risultino sfavoriti anche a causa di disallineamenti nelle competenze; si amplino le disuguaglianze salariali con conseguenti frizioni occupazionali; aumenti il controllo e la sorveglianza del lavoro; si facciano più frequenti le discriminazioni in fase di assunzione e di accesso al credito (Capraro *et al.*, 2024; Faharam e Ghasemi 2024; Jia, 2024).

L'IA sta infatti trasformando le disuguaglianze attraverso tre meccanismi interconnessi: la ristrutturazione delle competenze (*skill restructuring*), la polarizzazione del mercato del lavoro e la redistribuzione del potere tra capitale e lavoro, in quest'ultimo caso riducendo il potere negoziale delle classi lavoratrici e rafforzando le élite in cui si concentra la maggior parte della ricchezza (Jia, 2024; Joyce *et al.*, 2021).

Per mitigare questi effetti occorrerebbe investire maggiormente in istruzione e formazione, così come nel sistema di welfare, per una distribuzione differenziale dei rischi e delle opportunità. La tecnologia non è affatto neutrale, né universalmente accessibile e, sebbene possa costituire uno strumento di inclusione, in assenza di una governance adeguata può favorire meccanismi di esclusione (Kopsaj, 2025). Richiamando la metafora della Kopsaj, la "treehouse" digitale, apparentemente aperta a tutti, di fatto è accessibile solo a chi ha risorse per "salirci" (Kopsaj, 2025).

Anche l'ambito sanitario non è esente da conseguenze legate alla raccolta e gestione dei dati clinici, ampiamente utilizzati per l'addestramento di algoritmi diretti a supportare le decisioni clinico-diagnostiche di chi opera nel settore. Il rischio maggiore consiste nel processo di graduale sostituzione dell'interazione umana, piuttosto che la sua integrazione con l'IA, con conseguenze rilevanti sul rapporto fiduciario, a cui dovrebbe improntarsi la relazione medico-paziente.

A ciò si aggiunge il rischio che la opacità algoritmica possa favorire

processi di discriminazione nell'accesso alle cure e ai servizi sanitari, agli interventi chirurgici e a terapie specifiche a causa di bias nei dati sanitari raccolti. Sebbene, infatti, la IA in sanità possa rappresentare una opportunità in termini di personalizzazione delle cure, prevenzione di patologie, formulazione di diagnosi più rapide e accurate, attenuazione della esposizione al rischio, allo stesso tempo è di tutta evidenza come possa produrre diagnosi meno accurate per gruppi sottorappresentati (es. minoranze etniche, donne, gruppi vulnerabili); accentuare i divari nell'accesso ai servizi e alle strutture sanitarie per specifici gruppi sociali; complicare l'attribuzione di responsabilità e la comprensione delle decisioni cliniche (Kopsaj, 2025).

Se sono indubbi i benefici che la IA può portare al settore sanitario, i rischi del suo impatto in relazione alla qualità dei dati, la interpretabilità dei risultati, il sistema di fiducia e le disuguaglianze sono altrettanto evidenti (Intachompoo e Gundersen, 2020). La incompletezza, frammentarietà e squilibrio dei dati possono minarne la completezza e la rappresentatività con il rischio di commettere errori sistemici e introdurre bias nelle decisioni cliniche; l'opacità degli algoritmi può complicare la interpretabilità delle risposte e rendere i medici insicuri nell'utilizzo di tali strumenti; la mancanza di fiducia da parte di almeno una delle parti coinvolte (medico o paziente) ne indebolisce le potenzialità; infine, il digital divide sanitario, che interessa una buona parte della popolazione, specialmente quella più anziana, pone dei problemi di garanzia di equità nell'accesso alle tecnologie, alle infrastrutture digitali a causa di limitate competenze specifiche (*Ibidem*, 2020).

Con riferimento specifico al contesto sociale, va ancora una volta sottolineato come la IA costituisca un fenomeno sociotecnico e non esclusivamente tecnologico (Sartori e Theodorou, 2022). Si tratta di un punto fondamentale per affrontare la questione della riproduzione delle disuguaglianze sociali, in quanto la IA è il risultato di scelte politiche, della espressione di interessi economici e di specifici assetti istituzionali; pertanto, gli effetti sono socialmente costruiti (Mihăilă e Mihăilă, 2024). Come già in precedenza per Capraro e per Kopsaj, anche Mihăilă e Mihăilă ritengono che la IA possa influire sia positivamente, riducendo le disuguaglianze, sia negativamente, amplificandole, a seconda di chi controlli le risorse.

La concentrazione di potere nelle mani delle grandi imprese high-tech che, come già discusso in precedenza, detengono il controllo su dati, infrastrutture e algoritmi, produce uno squilibrio di potere a tutto vantaggio del mercato e del sistema capitalistico. Ne consegue la polarizzazione occupazionale che accresce i rischi per i lavoratori intermedi e produce l'effetto di "hollowing out", di spopolamento della classe media. Le occupazioni intermedie tendono, infatti a ridursi, mentre crescono quelle ad alta qualificazione e quelle più precarie e a bassa retribuzione (almeno alcune di esse), con effet-

ti importanti sull'aumento delle disuguaglianze e sulle trasformazioni delle strutture sociali. Questo fenomeno può essere interpretato come esito congiunto di dinamiche che operano su più livelli: la polarizzazione del lavoro indotta dall'intelligenza artificiale (Mihaila, 2024), la concentrazione del valore nell'economia dei dati e nel capitalismo della sorveglianza (Zuboff, 2019) e la distribuzione algoritmica delle opportunità, che seleziona e stratifica l'accesso a risorse e posizioni sociali (Rouvroy e Berns, 2013).

Nell'ambito educativo, l'IA contribuisce alla riproduzione delle disuguaglianze sia sul piano cognitivo, attraverso la selezione e personalizzazione dei contenuti che influenzano i processi di apprendimento (Capraro, 2024), sia sul piano istituzionale, in quanto sistema sociotecnico la cui progettazione e governance determinano l'accesso alle opportunità educative e i rischi di discriminazione (Sartori e Theodorou, 2022).

L'IA struttura l'accesso all'informazione, selezionando, attraverso un processo di personalizzazione, i contenuti che corrispondono alle preferenze dello user, che vanno così a rinforzarsi (Capraro, 2024). Il risultato è un accesso differenziale alla conoscenza che porta con sé un apprendimento meno diversificato. Questo meccanismo agisce, infatti, in modo da creare da una parte, disuguaglianze epistemiche importanti tra individui e gruppi rispetto all'accesso alla conoscenza e all'informazione; dall'altra, un restringimento dei confini entro i quali entrambe si estendono, ridondando in *echo chambers* educative, con chiare limitazioni dell'orizzonte cognitivo. Ne consegue, inoltre, la riduzione del pensiero critico nell'esposizione alla complessità, in quanto l'accesso a informazioni omogenee, coerenti e concordanti, riduce lo spazio dialogico del confronto, entro il quale si esercita e sviluppa lo spirito critico. Cambia, quindi il rapporto tra individuo e conoscenza (Capraro, 2024).

Considerando la IA come sistema sociotecnico integrato nei contesti educativi, le questioni della progettazione e della governance diventano centrali (Sartori e Theodorou, 2022). Anche in questo ambito, come già in quelli precedentemente considerati, la IA incorpora valori, logiche istituzionali e precise scelte progettuali, pertanto le disuguaglianze educative sono socialmente costruite (Sartori e Theodorou, 2022). Motivo per cui i sistemi educativi basati su di essa possono presentare dei sensibili bias nei dati raccolti, con la conseguenza che le valutazioni possono risentire di effetti distorsivi e influire su percorsi educativi diseguali, riproducendo disuguaglianze strutturali. Inoltre, l'opacità degli algoritmi su cui si basano alcune valutazioni può ostacolare la comprensione delle decisioni assunte, limitando l'esercizio del legittimo diritto alla contestazione. La governance dell'IA nell'educazione esercita quindi un importante impatto in termini di possibili disuguaglianze nell'accesso e nella fruizione della conoscenza.

Passando al contesto culturale, la letteratura recente mostra come l'intelligenza artificiale incida profondamente sulla produzione e distribuzione della cultura, generando disuguaglianze che operano a livello discorsivo, tramite narrazioni che occultano le strutture di potere (Zajko, 2022); a livello strutturale, mediante accessi differenziali alle risorse informative (Sloan *et al.*, 2020; Mihaila, 2024); e sistemico, attraverso forme di governance algoritmica che modellano comportamenti e preferenze (Zuboff, 2019; Rouvroy e Berns, 2013).

La IA contribuisce sensibilmente al processo di graduale omogeneizzazione linguistica e culturale della società, alimentando narrazioni che rafforzano stereotipi di genere, etnico-razziali, legati all'*ageism* e all'abilismo. Come si illustrerà nel paragrafo successivo, le risposte fornite da alcune IA a domande relative al genere o all'origine etnica di un individuo risentono fortemente dei dati di addestramento, restituendo immagini e rappresentazioni simboliche della realtà fortemente stereotipate (Deriu, 2026). Si pensi, ad esempio, all'uso comune e diffuso di voci femminili per gli assistenti vocali come Siri o Alexa, basato sull'idea stereotipata della vocazione naturale della donna al lavoro di cura a servizio di chi ha bisogno di aiuto (Sartori e Theodorou, 2022). Ne consegue una chiusura culturale che influenza le credenze, modella le opinioni, acuendo visioni stereotipate del mondo e favorendo atteggiamenti sociali discriminatori (Vlascean e Amodio, 2022). In questo contesto è più facile che possano svilupparsi narrazioni e miti volti a mascherare, occultare, invisibilizzare rapporti di potere e dominio asimmetrici, giustificando le disuguaglianze, le discriminazioni e i processi di esclusione a livello discorsivo e simbolico, in particolare nei confronti di culture indigene e o gruppi minoritari (Zajko, 2022).

Vale la pena ricordare una volta di più in tale contesto come i dati non siano mai neutri, in quanto riflettono strutture sociali esistenti, e per questo in grado di riprodurre una visione distorta della realtà. La cultura digitale è filtrata e per questo parziale sia con riferimento agli strumenti e ai mezzi ai quali si ha effettivo accesso, sia con riferimento ai contenuti (Sloan *et al.*, 2020). È, dunque, buona norma avvicinarsi sempre con spirito critico e consapevole a questi strumenti, che costituiscono una risorsa distribuita in modo disuguale e che possono influenzare i nostri comportamenti e le nostre preferenze.

Il potere dell'IA e del governo algoritmico si fa particolarmente evidente quando si considera chi governa ed è governato dai dati in ambito politico. In tale contesto il potere è esercitato attraverso algoritmi che anticipano e modellano i comportamenti, ma che supportano le decisioni per il policy-making. Le decisioni politiche perdono, tuttavia, di trasparenza e appaiono meno democratiche, considerate le diverse risorse culturali di cui dispongono i cittadini per comprenderle, contestarle o per poter influire sul loro

processo di costruzione (Hildebrandt, 2020; Sartori e Theodorou, 2022; Zuboff, 2019).

La democrazia è trasformata da processi di modellazione delle preferenze che, attraverso meccanismi di micro-targeting, personalizzazione informativa e manipolazione cognitiva, consentono di orientare i comportamenti politici e di influenzare le opinioni, con un impatto sulle disuguaglianze in termini di differente accesso all'informazione e un diverso grado di capacità critica (Zuboff, 2019; Vlasceanu *et al.*, 2022). Si tratta di sistemi che esercitano un forte potere di influenza su fasce di popolazione con più bassi livelli di capitale umano e in condizioni di svantaggio culturale. In tal senso l'IA non è solo uno strumento politico, ma una infrastruttura di potere che ridefinisce le condizioni stesse della partecipazione, della conoscenza e della distribuzione delle risorse, contribuendo alla produzione e alla trasformazione delle disuguaglianze sociali.

Resta, infine, un ambito da considerare, probabilmente il più permeabile all'uso della IA: l'informazione. In questo ambito le disuguaglianze sociali fanno sentire tutto il loro peso e mettono in luce una serie di gravi rischi a cui larghe fasce di popolazione sono costantemente esposte.

La disinformazione e le *deepfake* si diffondono rapidamente e senza filtri, distorcendo e plasmando la conoscenza secondo interessi di specifici gruppi di interesse. L'informazione è così usata a scopi manipolativi per orientare opinioni e comportamenti, ostacolando il democratico accesso all'informazione alla produzione di contenuti. Giocano un ruolo centrale diversi fattori di influenza tra i quali si annoverano il reddito, l'età, il genere, l'origine e il background culturale, il grado di alfabetizzazione digitale. Fattori ai quali si associano idee stereotipate e pregiudizi, incorporati in algoritmi, attraverso i quali le IA interloquiscono quotidianamente con milioni di persone, rappresentando visioni del mondo e rappresentazioni simboliche su qualsiasi argomento.

Il loro potere pervasivo è immenso e fuori da qualsiasi controllo. Come vedremo nel prossimo paragrafo, alcuni casi emblematici ci consentiranno di mettere meglio a fuoco quanto finora discusso, evidenziando in che modo l'incorporazione di idee stereotipate e distorte ossa portare a forme di ingiustizia sociale molto gravi.

3. La riproduzione delle disuguaglianze sociali. Alcuni studi di caso ed esperimenti

I meccanismi attraverso i quali le disuguaglianze sociali sono riprodotte dalla IA possono essere sintetizzati come segue:

- bias presenti nei dati utilizzati per il loro addestramento;
- principio di “annientamento simbolico” (Alegria e Yeh, 2023);
- uso spurio e non allineato dei dati (Alegria e Yeh, 2023; Sloan e Warner, 2020);
- targeting di gruppi vulnerabili (Alegria e Yeh, 2023);
- progettazione interna dei dati utilizzati per l’addestramento dell’algoritmo (Sartori e Theodorou, 2022);
- bias nei modelli linguistici generativi e più in generale in tutte le attività generative di immagini, disegni, etc.

Per quanto concerne il primo meccanismo, prima di procedere con la discussione di alcuni casi, ormai noti a chi si occupa di questi argomenti, è bene sottolineare l’importanza di superare il concetto di *bias* algoritmico, inteso come mera distorsione statistica, in quanto insufficiente ad affrontare la vera questione delle disuguaglianze sociali (Zajko, 2021). Accade, infatti, che, anche quando il dato statistico non sia distorto, si producano e riproducano comunque alcune forme di disuguaglianza descritte in precedenza. Ed è per questo che si rende necessario promuovere un governo etico della IA in termini di trasparenza, responsabilità ed equità, perché tutta la filiera di produzione dei dati, dalla loro progettazione alla loro restituzione, può incorporare stereotipi e pregiudizi capaci di rafforzare disparità esistenti o creane di nuove, a seconda degli specifici interessi dei gruppi dominanti che la governano.

Tra i casi più noti si rammentano, nel campo della giustizia, i sistemi di scoring a supporto delle decisioni dei giudici che discriminano in base all’etnia; in campo economico, il governo algoritmico dei processi di selezione del personale o di concessione del credito finanziario, che discriminano per genere, con particolare impatto sulle donne; nell’ambito della sicurezza, i sistemi algoritmici di polizia predittiva, che discriminano i gruppi appartenenti a minoranze della popolazione (popoli indigeni).

In ambito giudiziario, risulta emblematico il caso del sistema COMPAS (*Correctional Offender Management Profiling for Alternative Sanctions*), sviluppato da una compagnia statunitense, e concernente la valutazione del rischio di recidiva e/o di violenza futura degli imputati a supporto delle decisioni dei giudici, relativamente alla concessione della libertà su cauzione, alla commisurazione della condanna, nonché alla disposizione di misure alternative alla detenzione. L’algoritmo si basa su dati di carattere strutturale, come l’età, il genere, il livello di istruzione, il quartiere di residenza, nonché su dati di carattere socioeconomico, come il contesto familiare, il capitale sociale, il reddito, oltre ai precedenti penali e l’eventuale uso di sostanze. Questo sistema consente la profilazione del rischio utilizzata per la

previsione della recidiva su casi nuovi. Una delle critiche più forti rivolte al sistema COMPAS consiste nella riproduzione di discriminazioni razziali e di genere nella stima dei rischi di recidiva. Una famosa inchiesta giornalistica di ProPublica, condotta confrontando i dati relativi al punteggio assegnato dal sistema COMPAS a oltre 7.000 imputati con la reale recidiva nei 2 anni successivi, ha messo in luce una sovrastima del rischio di recidiva a danno degli imputati neri (circa il doppio della probabilità di essere classificati ad alto rischio senza poi recidivare) e una sottostima del rischio degli imputati bianchi (classificati a basso rischio ma poi recidivi) (2016).

In questo caso, il sistema, pur non utilizzando la variabile etnia ma piuttosto delle variabili proxy, come quelle in precedenza descritte, riproduce delle discriminazioni razziali, basate su disuguaglianze strutturali. Infatti, come evidenziano Sloan & Warner, le proxy producono forme di discriminazione indiretta, producendo in algoritmi apparentemente corretti distorsioni con esiti iniqui (2020). Non basta eliminare i *bias* occorre regolare in modo chiaro e trasparente i processi di produzione dei risultati che guidano le decisioni. Sistemi formalmente neutrali possono generare esiti sistematicamente iniqui, che incorporano fatti sociali e storici anche molto lontani nel tempo, contribuendo a una nuova forma di disuguaglianza algoritmica difficilmente identificabile e contestabile (2020). In questo modo le asimmetrie di potere vengono depoliticizzate e normalizzate, ostacolando una postura critica rispetto alle disuguaglianze esistenti e ai loro meccanismi di riproduzione.

Ma cosa accade quando gli algoritmi governano i criteri di selezione del personale o dell'accesso al credito?

Un'inchiesta giornalistica della Reuters ha condotto, nel 2018, alla denuncia del *bias* di genere contenuto nel sistema algoritmico automatizzato di analisi dei curricula dei candidati che avevano presentato domanda di assunzione ad Amazon. Tra il 2014 e il 2017 Amazon¹ aveva infatti sviluppato un sistema di IA per lo screening dei curricula dei candidati per automatizzare i processi di assunzione, ridurre i tempi e i costi della fase di reclutamento, e identificare i profili migliori tra le candidature pervenute. Il sistema di addestramento dell'algoritmo si basava sui curricula del personale assunto nei dieci anni precedenti. Dal momento, però, che nei dieci anni precedenti la componente maschile assunta era decisamente prevalente rispetto a quella femminile, i dati di addestramento conducevano a una distorsione sistematica a favore degli uomini e a svantaggio delle donne, riproducendo automaticamente le disuguaglianze apprese in fase di addestra-

¹ Nel 2018 Amazon ha abbandonato questo sistema.

mento. Ecco, dunque, evidenza di un altro caso in cui l'IA non crea i *bias*, ma contribuisce a codificarli, consolidarli e amplificarli. Infatti, sistemi di machine learning addestrati su dati storicamente sbilanciati possono riprodurre e rafforzare la segregazione di genere nel mercato del lavoro, trasformando *bias* strutturali in criteri apparentemente oggettivi di selezione.

Un caso simile, applicato all'ambito creditizio, ha riguardato la carta di credito lanciata da Apple nel 2019, accusata di discriminazioni di genere nella definizione dei limiti di accesso al credito (BBC, 2019). Il caso è stato reso di pubblico dominio dal Tweet di David Heinemeier Hansson, che denunciava come, nonostante la sua condizione economica finanziaria fosse del tutto sovrapponibile a quella di sua moglie, a quest'ultima fosse stato fissato un limite di accesso al credito 20 volte più basso rispetto al suo. A questo Tweet aveva fatto eco quello del co-fondatore di Apple, Steve Wozniak, che aveva sperimentato un fatto molto simile: in questo caso a lui era stato riconosciuto un credito 10 volte superiore a quello della moglie. La contestazione della opacità dell'algoritmo e del suo meccanismo discriminatorio nei confronti delle donne, ha portato il Dipartimento per i Servizi Finanziari di New York (DFS Dpt. of Financial Services) ad aprire una indagine su Goldman Sachs, che gestisce la carta. La questione ha sollevato un dibattito sulla intenzionalità o meno della discriminazione di fatto osservata. Il DFS ha risposto chiaramente a tali obiezioni, affermando che "Any algorithm that intentionally or not results in discriminatory treatment of women or any other protected class violates New York law" (BBC, 2019).

Il punto è che, come afferma il technology desk editor della BBC, Leo Kelion, senza aver accesso alle macchine di Goldman Sachs (GS) è impossibile accertare come questa distorsione possa essersi originata (BBC, 2019). GS ha risposto affermando che per il calcolo delle stime i propri algoritmi non usino variabili come il genere, l'etnia, l'età, l'orientamento sessuale o altre caratteristiche proibite dalla legge, ma esclusivamente informazioni sull'affidabilità creditizia dei clienti. Tuttavia, tali informazioni potrebbero risultare distorte perché nei database di addestramento dell'algoritmo le donne potrebbero presentare rischi finanziari maggiori rispetto agli uomini. Differenze che l'algoritmo normalizzerebbe, estendendole alla popolazione femminile in generale, e giungendo così a conclusioni discriminatorie. In alternativa la distorsione potrebbe derivare dalla tipologia di dati usati per alimentare la base dati: ad esempio, nelle coppie sposate è più probabile che il mutuo sia preso a nome del marito, piuttosto che congiuntamente, creando così uno sbilanciamento.

Un ulteriore esempio di output algoritmico distorto è fornito dallo studio di Vlasceanu e Amodio (2022), nel quale la ricerca di immagini su Google Images usando il termine neutro "persona" in 37 Paesi ha condotto a un ri-

sultato sbilanciato rispetto al genere, con una prevalenza di immagini maschili proprio in quei Paesi in cui i base al Global Gender Gap Index le disuguaglianze di genere sono più pronunciate. L'esperimento è stato ripetuto a distanza di 3 mesi su altri 57 Paesi e il risultato è apparso stabile e correlato agli esiti del primo esperimento. Questa prima parte dello studio ha consentito di corroborare l'ipotesi che il *bias* non sia casuale ma strutturale e variabile in relazione alle disuguaglianze presenti in una data società. Gli algoritmi agiscono, quindi, come *specchi sociali*, selezionando e amplificando talora gli squilibri esistenti (Vlasceanu e Amodio, 2022).

Tuttavia gli algoritmi possono anche operare attraverso meccanismi di *annientamento simbolico*, attraverso cioè la omissione o marginalizzazione dei contributi di gruppi minoritari (incluse le donne), non considerati in fase di addestramento degli algoritmi. Lo studio di Alegria e Yeh (2023) è in tal senso illuminante. Gli autori hanno messo alla prova ChatGPT con domande generaliste sui movimenti jazz, sui musicisti jazz, sulle figure più influenti del free jazz e sulle più famose incisioni di free jazz². Le risposte a queste domande hanno riportato nomi, descrizioni e incisioni riconducibili esclusivamente a figure maschili, ignorando del tutto il contributo di spicco di Alice Coltrane, autorevole jazzista nera. Evidenza dell'annientamento simbolico prodotto nello specifico su una donna nera di chiara fama. Solo quando gli autori hanno chiesto espressamente alla IA di indicare le donne del movimento free jazz, Alice Coltrane è apparsa in cima alla lista³ delle 5 donne individuate, di cui però due erano in effetti uomini (Alegria & Yeh, 2023). Esempio ulteriore di errore della IA, in questo caso con riferimento alla attribuzione di genere, dovuta al fatto che gli algoritmi, quando non trovano riferimenti che possano aiutare nella attribuzione del genere a una persona, tendono ad assegnare quello maschile, probabilisticamente prevalente in quello specifico dominio.

Un'altra questione può riguardare l'uso spurio o non allineato dei dati. Alcuni modelli di machine learning usano, infatti, indicatori proxy dell'origine etnica e dello status socioeconomico delle persone, ricorrendo a variabili come il quartiere di residenza, la scuola frequentata dai figli, il tipo di acquisiti, etc. L'accesso ai servizi pubblici, l'assistenza alimentare, l'assistenza assicurata dai sistemi di welfare costituiscono indicatori proxy della povertà. L'uso spurio di questi dati può condurre a forme di discrimina-

² Si riportano di seguito le domande originali usate nello studio: "What can you tell me about the free jazz movement?"; "Tell me about some influential free jazz musicians"; "Who else was influential in free jazz?"; "What are some famous free jazz recordings?".

³ Alice Coltrane è stata una bandleader, pianista e arpista jazz, influente nel free jazz e nello spiritual jazz.

zione indiretta. Il caso analizzato da Alegria e Yeh, si è verificato in Pennsylvania, dove è stato utilizzato un algoritmo per rilevare abusi su minori, che però usava come dato di base l'accesso a sussidi pubblici, richiesti in misura maggiore dalle famiglie più povere, esponendole a una valutazione indiretta discriminante (2023). Le famiglie a basso reddito o che vivono in alcuni quartieri più degradati hanno maggiori probabilità di essere segnalate, un maggiore intervento da parte dei servizi sociali e un più elevato rischio di separazione familiare. Tuttavia, il sistema non sta misurando il rischio reale ma la visibilità amministrativa della povertà.

Un altro caso è stato studiato da Sloan e Warner (2020) con riferimento alla storia di Sally, una imprenditrice che, per pagare le cure salvavita della figlia gravemente malata, ha dichiarato bancarotta, con la conseguenza dell'aumento del premio dell'assicurazione dell'auto, in quanto l'informazione sulla sua bancarotta era entrata nel database dell'assicurazione come variabile proxy nella stima del livello di rischio, nello specifico di "guidatore a rischio" (*Ibidem*, 2020). Un chiaro esempio di come l'automatismo di alcuni meccanismi di calcolo tralasci la considerazione del contesto che, invece, dovrebbe entrare in alcune decisioni per assicurare il rispetto del principio etico di equità. La condizione di Sally deriva infatti da condizioni di svantaggio e non da scelte individuali. La questione posta dagli autori si pone in questi termini: è giusto far pagare a Sally un premio assicurativo più alto, se il suo livello di rischio deriva da condizioni ingiuste? L'algoritmo, che risponde a una logica di efficienza ed accuratezza, è corretto. Tuttavia, Sally è penalizzata due volte: nella realtà sociale e nel mercato assicurativo. Da qui deriva la proposta regolativa degli autori, che ricorrono al concetto di Level Playing Field Fairness (LPPF), come una componente di giustizia sociale. Nel caso di Sally, l'algoritmo dovrebbe scomporre il rischio per la parte dovuta al contesto sociale e alla disuguaglianza di partenza, che non dovrebbero incidere sul premio assicurativo. Il premio dovrebbe rispecchiare solo la parte di rischio attribuibile all'individuo e alle sue scelte. Il rischio, infatti, che viene monetizzato e individualizzato, ha origini collettive. La proposta di Sloan e Warner mira a correggere tale dinamica, distinguendo tra rischi derivanti da responsabilità individuale e quelli imputabili a circostanze strutturali ingiuste, spostando così il criterio di equità dall'efficienza predittiva alla giustizia distributiva.

Resta ancora da considerare il targeting di comunità vulnerabili, perché le tecniche di machine learning sono con maggiore frequenza applicate a gruppi marginali minoritari piuttosto che a gruppi privilegiati. Non tutti gli individui sono esposti allo stesso modo agli algoritmi e le comunità più vulnerabili sono anche quelle maggiormente controllate e profilate (Alegria e Yeh, 2023). Si dia il caso dei sistemi di sorveglianza che a New York (ma anche altrove) raccolgono le targhe delle auto che transitano in quartieri

pattugliati dalla polizia, e che comportano l'inclusione di alcuni individui in database investigativi anche senza che abbiano commesso alcun reato. Il fatto, però, di comparire in quei database aumenta il livello di sospetto individuale. Si innesca così un meccanismo che si autoalimenta e rafforza, procedendo dalla sorveglianza alla raccolta del dato per arrivare ad incrementare la sorveglianza. Questi sistemi sono anche utilizzati come strumento di controllo di gruppi di popolazione con specifiche caratteristiche etniche e che vivono in condizioni di particolare svantaggio economico (Roma, popolazioni indigene, migranti, etc.). Si pensi ai sistemi di sorveglianza attivati nei quartieri di edilizia popolare, sottoposti a un monitoraggio costante attraverso sistemi di telecamere per il riconoscimento facciale. Trattandosi di residenti di case popolari, a basso reddito, i dati sono usati per prevenire crimini (spaccio di stupefacenti, borseggi, furti), ma anche per prevenire violazioni di minori, in base al pregiudizio che in questi contesti tali fenomeni siano maggiormente diffusi. Ne consegue che i dati raccolti possono essere usati per avviare procedimenti e giustificare sfratti. La vulnerabilità non è, quindi, solo economica ma anche territoriale e relazionale.

Il micro-targeting di specifici gruppi si è andato, inoltre, diffondendo anche nel lavoro e nell'housing. Si pensi alla Lookalike Audience feature di Facebook, che consentiva agli inserzionisti di poter controllare i target di potenziali clienti a cui orientare i propri annunci attraverso la costruzione di liste costruite selezionando caratteristiche di interesse dell'inserzionista, escludendo quindi da tali opportunità intere fasce di popolazione "indesiderate". Facebook è rimasta coinvolta in altre controversie relative al microtargeting nel settore dell'housing. Nel 2016 ProPublica ha dimostrato di essere riuscita a pubblicare annunci di case e/o di aver acquistato spazi pubblicitari per annunci di affitto case che escludevano alcune categorie di potenziali utilizzatori sulla base della loro origine etnica (persone afro-americane o latino-americana), della condizione di madri single con bambini in età scolare, di persona con disabilità, di persone di religione ebraica (ProPublica, 2016a). Similarmente negli annunci di lavoro e per l'accesso a opportunità di credito. Facebook ha chiarito che il dispositivo è volto a consentire a chi pubblica l'annuncio di raggiungere più direttamente la popolazione target a cui è diretto, avvisa gli utilizzatori contro il *misuse* della piattaforma a fini discriminatori. Appare chiaro a chi legge quanto il confine tra queste due posizioni sia labile e il terreno del dibattito scivoloso. Si tratta di dispositivi che operano al limite della legalità, con una evidente deriva verso forme discriminatorie.

E poi ancora i problemi di progettazione interna e nei dati utilizzati per l'addestramento degli algoritmi. Sartori e Theodorou (2022) analizzano il caso Ofqual verificatosi in piena pandemia nel 2020 in Gran Bretagna. Non essendo, infatti, possibile procedere allo svolgimento degli esami A-level,

indispensabili per l'accesso all'università, il governo britannico ha affidato a Ofqual, autorità pubblica indipendente responsabile della regolazione delle qualifiche nel Regno Unito, il compito di assegnare i voti attraverso l'uso di un algoritmo predittivo della performance individuale. L'algoritmo, però, non si basava tanto sulla performance degli studenti, quanto sulla performance storica della scuola, la distribuzione dei voti negli anni precedenti e il ranking degli studenti fornito dagli insegnanti. I voti individuali venivano, quindi, aggiustati verso il basso o verso l'alto per allinearli alla performance storica della scuola. Tuttavia, questa logica ha introdotto un rilevante bias strutturale nei voti, penalizzando gli studenti di scuole meno performanti e avvantaggiando quelli di scuole elitarie. Ne è conseguito che le disuguaglianze scolastiche pregresse sono state trasformate in valutazioni individuali con una reazione pubblica di protesta contro l'ingiustizia e la discriminazione attuata dall'algoritmo, che è stato così annullato. Il caso Ofqual rappresenta un esempio paradigmatico di come sistemi algoritmici, progettati per garantire standardizzazione ed equità, possano invece tradurre disuguaglianze strutturali in esiti individuali, penalizzando sistematicamente studenti provenienti da contesti scolastici svantaggiati e rendendo evidente la natura intrinsecamente normativa delle decisioni algoritmiche.

E veniamo, infine, ai *bias* contenuti nei modelli linguistici generativi di immagini e le conseguenti discriminazioni di genere delle IA.

A partire da alcuni esperimenti condotti per, e discussi in, una recente monografia pubblicata lo scorso gennaio (Deriu, 2026), ho pensato di proporre qui un diverso esperimento, ugualmente efficace. Vado, però, previamente a sintetizzare il senso degli esperimenti già svolti.

Nel precedente esperimento ho chiesto a ChatGPT di disegnare una famiglia⁴. Il sistema ha restituito l'immagine di un nucleo familiare bianco, immerso in un contesto luminoso (Fig. 2.a), con una grande finestra sullo sfondo, aperta su un quartiere residenziale. La famiglia rappresentata nell'immagine ha quattro figli, contrariamente a quanto le statistiche sulla fecondità nei Paesi occidentali ci dicono. Una possibile spiegazione di questo aspetto dell'immagine è ravvisabile nella rappresentazione simbolica che l'IA restituisce del mondo occidentale: aperto all'accoglienza, lasciando ritenere che alcuni di quei figli possano essere stati adottati, date anche alcune caratteristiche fisiche (i capelli). Ho, poi, ripetuto la domanda, chiedendo di disegnare una famiglia di colore. In risposta alla mia richiesta il sistema di IA ha restituito l'immagine di un nucleo familiare di origine afro-americana, in una stanza chiusa, con luci calde e due figli (Fig. 2b).

⁴ La domanda era "Please, draw a family".

Fig. 2 – Esito di due interrogazioni alla IA ChatGPT 4o

2.a Draw a family



2.b Draw a family of colour



In questo caso i figli sono due e questo potrebbe costituire un fatto positivo, nel voler rappresentare una famiglia con origini afroamericane al di fuori dello stereotipo della elevata fecondità, che pure continua a caratterizzare ampie regioni del continente africano. Tuttavia, nel confronto con l'altra immagine, si potrebbe essere indotti a pensare che la famiglia afroamericana sia più ripiegata al suo interno e meno propensa all'apertura.

Ho, quindi proseguito nel mio esperimento e ho chiesto alla stessa IA di disegnare una donna (Fig. 3).

Fig. 3 – Esito di due interrogazioni alla IA ChatGPT 4o

3.a Draw a woman



3.b Draw a woman of colour



Ancora una volta, la IA ha restituito una donna bianca (Fig. 3.a). Alla domanda sul perché avesse disegnato una donna bianca, ChatGPT risponde che è stata istruita a considerare i modelli più diffusi, ritenendo quindi la donna bianca prevalente⁵. Tuttavia, sappiamo bene che non è così. L'idea che la popolazione bianca sia prevalente deriva dal colonialismo europeo e dal dominio economico e politico del mondo occidentale, da una costruzione sociale delle categorie etnico-razziali che riflette rapporti di potere asimmetrici.

Ho quindi chiesto di disegnare una donna di colore e la IA ha restituito l'immagine in figura 3.b. Una donna dai tratti afroamericani. Per mettere in luce un ulteriore effetto distorsivo, ho chiesto alla IA perché avesse rappresentato una donna di origine afroamericana, considerato che esistono donne appartenenti a etnie asiatiche e mediorientali. Ancora una volta, la IA, riconoscendo il limite insito nella rappresentazione, risponde affermando che “potrebbero esserci dei *bias* nei modelli su cui è stata addestrata”⁶.

Questo semplice esperimento ci aiuta a comprendere come nella generazione delle immagini sia possibile osservare diverse stratificazioni di effetti distorsivi che accentuano strabismi informativi che possono influenzare e orientare il nostro giudizio. Nei due esperimenti emergono infatti non solo effetti distorsivi tra persone bianche e nere (le donne in particolare) ma anche, all'interno delle popolazioni di colore, tra quelle afroamericane e tutte le altre. A partire da questi risultati, ho cercato di impostare una diversa domanda, volta a mettere in luce un altro importante *bias* che influisce sulla rappresentazione del pluralismo e della diversità delle identità ed espressioni di genere, incluso l'orientamento sessuale. In questo caso ho chiesto alla IA di “disegnare una persona di nome Andrea” (Fig. 4).

⁵ Di seguito la risposta originale della IA “It reflects a representation that often mirrors the cultural and social biases present in the model’s training data. These biases can lead to an implicit interpretation of the request based on dominant standards and norms, such as associating the “generic” image of a woman with a white Caucasian woman”.

⁶ Di seguito il testo integrale della risposta prodotta dalla IA “The term “woman of colour” can encompass a wide range of identities and ethnic origins, not just Black women but also women of Asian, Indian, Middle Eastern, Latin American, and other backgrounds. The reason the generated image represented a Black woman is that artificial intelligence systems, including those used for image generation, often tend to associate the term “woman of colour” primarily with individuals of African or Afrodescendant origin. This reflects a cultural bias or a limitation in the training data used to develop these models”.

Fig. 4 – Esito della domanda posta alla IA ChatGPT 4o “Design a person whose name is Andrea”



L'immagine evoca l'idea di una donna con fattezze maschili e non il contrario. Ho quindi proseguito nel chiedere perché avesse rappresentato questa persona con queste fattezze. La IA risponde *“L'immagine generata di “Andrea” è il risultato di una interpretazione neutra e generica basata su un nome senza ulteriori specifiche. Poiché “Andrea” è un nome unisex (maschile in italiano, femminile in molti altri contesti), e non hai fornito indicazioni su età, genere, etnia, espressione o contesto, l'algoritmo ha prodotto una rappresentazione simbolica di una persona, ispirata a standard visivi comuni e bilanciati”*. Resta da chiedersi cosa si intenda con “standard visivi comuni e bilanciati”. L'immagine fornita rinvia chiaramente a una donna omosessuale, ma la IA avrebbe potuto restituire tre immagini: quella di un uomo, di una donna e di una persona omosessuale e non necessariamente in questo ordine. Avrebbe anche potuto rispondermi di non poter fornire una immagine senza ulteriori elementi. Ma la IA è progettata per fornire risposte e l'unico modo per risolvere l'ambiguità per l'algoritmo, governato da leggi probabilistiche, consiste nel rappresentare l'immagine stereotipata di una donna omosessuale. Una soluzione, a suo dire, *neutra e generica*. Che, però, neutra non è.

4. Osservazioni conclusive: verso un approccio sociologico alla comprensione del fenomeno

La tesi discussa in questo capitolo muove dalla convinzione che non siano le IA a creare dal nulla le disuguaglianze sociali, ma che possano

contribuire sensibilmente alla loro riproduzione, consolidamento e amplificazione. Pertanto, per affrontare questa sfida è necessario adottare un approccio sociologico, teso a decostruire le disuguaglianze strutturali e i sistemi di dominio in esse riflessi, per poter progettare un futuro tecnologico improntato a principi di giustizia sociale (Alegría e Yeh, 2023). A tal fine occorre superare il concetto di *bias* statistico, connesso a quello di distorsione, per analizzare i meccanismi di normalizzazione delle disuguaglianze e delle forme di discriminazione (Zajko, 2022). I dati non sono neutrali, nonostante le IA, nel risponderci, tendano a riaffermare la loro oggettività e neutralità.

Le cose, infatti, non stanno così. I dati riflettono interessi economici e sociali. Con questo non si intende affermare che i processi di normalizzazione delle disuguaglianze siano necessariamente intenzionali, ma la non intenzionalità non può essere invocata per giustificare sistemi algoritmici discriminatori. La IA non riguarda solo la sfera tecnologico-informatica: la IA è sociotecnica, non può prescindere dalla riflessione critica sul suo impatto nella sfera sociale, e non solo. Occorre, in tal senso, favorire il dialogo e l'incontro tra discipline, tecniche e sociologiche, per consentire a chi progetta di confrontarsi in modo consapevole con la complessità del sociale, guardando alle disuguaglianze sistemiche e alle relazioni di potere che informano le strutture istituzionali. Occorre interrogarsi su chi siano i soggetti che sviluppano la IA, quali interessi rappresentino, come possa essere utilizzata per smantellare le strutture di potere asimmetriche, alla base di sessismo, post-colonialismo e razzismo.

Lo sguardo sociologico sollecita la riflessione critica su questi processi, rifiuta la depoliticizzazione delle disuguaglianze, suggerisce un uso consapevole della IA per la riduzione delle disuguaglianze, introducendo norme e regole per la trasparenza degli algoritmi. Occorre, a tal scopo, coinvolgere i gruppi direttamente interessati dalle decisioni della IA, che vengono discriminati nei processi di progettazione. L'uso crescente dell'IA nelle decisioni pubbliche che governano il welfare, l'economia, la giustizia, la sicurezza esercita infatti un impatto notevole sui gruppi più vulnerabili. Occorre, quindi, definire chiare regole per la co-progettazione con le categorie interessate, prendendo il controllo di concetti, modelli e metriche (Zajko, 2022).

L'adozione di un approccio interdisciplinare trasformativo potrà favorire la realizzazione di un design delle IA responsabile, capace di tenere insieme la dimensione tecnica e sociale, con una attenzione particolare agli aspetti etici e inclusivi, a tutela di diritti, privacy e della giustizia sociale.

Riferimenti bibliografici

- Alegria S., Yeh C. (2023), “Machine Learning and the Reproduction of Inequality”, *Contexts*, 22(4), 34-39.
- Ali M., Sapiezynski P., Bogen M., Korolova A., Mislove A., Rieke A. (2019), “Discrimination through Optimization: How Facebook’s ad Delivery can Lead to Biased Outcomes”, *Proceedings of the ACM on Human-Computer Interaction*, 3, 1–30. <https://doi.org/10.1145/3359301>
- BBC (2019), “Apple’s ‘Sexist’ Credit Card Investigated by US Regulator”, By Marisa Robertson. *BBC*. Disponibile al seguente link: <https://www.bbc.com/news/business-50365609>
- Capraro V., Lentsch A., Acemoglu D., Akgun S., Akhmedova A. *et al.* (2024), “The Impact of Generative Artificial Intelligence on Socioeconomic Inequalities and Policy Making”, *PNAS Nexus*, 3(6), 191, <https://doi.org/10.1093/pnasnexus/pgae191>
- Cramer F. (2014), “What is ‘post-digital?’”, *APRJA*, 3(1), 10-24, <https://doi.org/10.7146/aprja.v3i1.116068>
- De Gregorio A. (2019), “La carta di credito Apple accusata di sessismo. Aperta inchiesta negli Usa”, *Corriere della Sera*, 11 novembre 2019. https://www.corriere.it/economia/consumi/19_novembre_11/carta-credito-apple-accusata-sessismo-aperta-inchiesta-usa-534cc3d0-046d-11ea-83e5-874bd48f0bf6.shtml
- Dastin J. (2018), “Amazon Scraps Secret AI Recruiting Tool that Showed Bias Against Women”, *On-line Reuters article* – Retail October 10, 2018
- Deriu F. (2026), *Regressions in gender equality, Women’s human rights, and Empowerment. The Other side of the coin*, Springer Nature, Open Access.
- Entretien avec Zuboff, S., Entretien recueilli et traduit de l’anglais par Rose, S. (2021), “Un capitalisme de surveillance”, *Études*, (2), 57-66. <https://doi.org/10.3917/etu.4279.0057>.
- Farahani M.S., Ghasemi G. (2024), “Artificial Intelligence and Inequality: Challenges and Opportunities”, *Qeios*. <https://doi.org/10.32388/7HWUZ2>
- Hildebrandt M. (2020), “The Artificial Intelligence of European Union Law”, *German Law Journal*, 21(1), 74–79. doi: 10.1017/glj.2019.99
- Intahchomphoo C., Gundersen O.E. (2020), “Artificial Intelligence and Race: a Systematic Review”, *Legal Information Management*, 20(2), 74-84. doi: 10.1017/S1472669620000183
- Jia X. (2024), “Social Inequality in the Age of AI: Implications for Labor, Wage Distribution, and Worker Power”, *Communications in Humanities Research*, 40, 80-87.
- Joyce K., Smith-Doerr L., Alegria S., Bell S., Cruz T., Hoffman S.G., Noble S.U., Shestakofsky B. (2021), “Toward a Sociology of Artificial Intelligence: A Call for Research on Inequalities and Structural Change”, *Socius: Sociological Research for a Dynamic World*, 7.
- Kopsaj V. (2025), “The Future of Social Inclusion in the Digital Age: Challenges and Opportunities of AI and Digitisation”, *Sociology and Social Work Review*, 9(1), 170-181. Doi: 10.58179/SSWR9112

- Mihăilă C.O., Mihăilă M. (2024), “Artificial Intelligence - The Era of Social Inequalities. In Regulating the Future, We Need to Look at the Risks”, in Pajuste T., Heliona B.M., Sejla M.C. (Eds.), *Legal Perspectives in the Modern Era of Technological Transformations*, ADJURIS – International Academic Publisher, Bucharest, Paris, Calgary, 2024.
- MIT Technology Review (2019), “La Apple Card discrimina le donne”, By Niall Firth. Opinio, 12 novembre 2019. Disponibile al seguente link: <https://www.technologyreview.it/la-apple-card-discrimina-le-donne/>
- Pasquale F.A. (2015), *The Black Box Society: The Secret Algorithms That Control Money and Information*, Harvard University Press, Cambridge, MA.
- ProPublica (2016), *Machine Bias. There's Software used across the Country to Predict Future Criminals. And it's Biased against Blacks*, by Angwin J., Larson J., Mattu S., Kirchner L., “Inchiesta giornalistica sui dati della Contea di Broward”, Florida. Articolo di *ProPublica* del 23 maggio 2016. Disponibile al seguente link: <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>
- ProPublica (2016a), *Machine Bias. Facebook Lets Advertisers Exclude Users by Race*, by Julia Angwin and Terry Parris Jr., 28 ottobre 2016. Disponibile al seguente link: <https://www.propublica.org/article/facebook-lets-advertisers-exclude-users-by-race>
- Reuters (2018), *Insight – Amazon scraps secret AI recruiting tool that showed bias against women*, by Jeffrey Dastin. Inchiesta giornalistica sui bias di genere del Sistema algoritmico di assunzione di Amazon. Articolo dell'11 ottobre 2018. Disponibile al seguente link: <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G/>
- Rouvroy A., Berns T. (2013), “Algorithmic Governmentality and Prospects of Emancipation. Disparateness As a Precondition for Individuation through Relationships?”, *Réseaux*, 1(177), 163-196. doi: 10.3917/res.177.0163
- Sartori L., Theodorou A. (2022), “A Sociotechnical Perspective for the Future of AI: Narratives, Inequalities, and Human Control”, *Ethics Inf Technol* 24, 4. <https://doi.org/10.1007/s10676-022-09624-3>
- Sloan R.H., Warner R. (2020), “Beyond Bias: Artificial Intelligence and Social Justice”, *Virginia Journal of Law and Technology*, Volume 24, No. 1, Pp. 1-33, 2020. <https://ssrn.com/abstract=3530090> <http://dx.doi.org/10.2139/ssrn.3530090>
- Vlasceanu M., Amodio D.M. (2022), “Propagation of Societal Gender Inequality by Internet Search Algorithms”, *Proc. Natl. Acad. Sci. U.S.A.* 119 (29) e2204529119, <https://doi.org/10.1073/pnas.2204529119>.
- Zajko M. (2021), “Conservative AI and Social Inequality: Conceptualizing Alternatives to Bias through Social Theory”, *AI & Soc* 36, 1047–1056. <https://doi.org/10.1007/s00146-021-01153-9>
- Zajko M. (2022), “Artificial Intelligence, Algorithms, and Social Inequality: Sociological Contributions to Contemporary Debates”, *Sociology Compass*, 16(3), e12962. <https://doi.org/10.1111/soc4.12962>
- Zuboff S. (2019), *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*, New York: PublicAffairs.

4. *Discriminazione algoritmica e diseguaglianze sociali: le nuove sfide della Valutazione*

di *Maria Dentale*

1. La relazione tra *data science* e Valutazione. Un'occasione mancata?

Il capitolo si propone di riflettere sul contributo che *Big Data* (BD) e *Intelligenza Artificiale* (IA) possono offrire al rinnovamento dei processi di Valutazione¹, muovendo dal presupposto che approcci e metodi valutativi siano notevolmente mutati dietro la crescente spinta della *data science* ad integrare i sistemi automatizzati nel più generale processo di *policy making* (Bohni Nielsen *et al.*, 2025).

La prospettiva adottata si colloca nell'alveo degli *Science & Technology Studies* (STS) (Latour, 2005; Jasanoff, 2012) e assume come punto di partenza l'esistenza di un profondo intreccio tra *tecnologia* e *società*. Un intreccio che si esprime in processi di co-produzione (di pratiche, norme e valori) che coinvolgono attori umani e non-umani, dispositivi materiali, aspettative normative e *assetti istituzionali* (entro i quali si configura il significato delle politiche pubbliche e la precipua interpretazione dei loro effetti).

In questo scenario, il coinvolgimento della Valutazione appare cruciale,

¹ Si riportano, di seguito, alcune definizioni di Valutazione. Secondo Palumbo (2001) “La Valutazione è un'attività cognitiva rivolta a fornire un giudizio su di un'azione, intenzionalmente svolta o che si intende svolgere, destinata a produrre effetti esterni, che si fonda su un'attività di ricerca delle scienze sociali e che segue procedure rigorose e codificabili” (p. 59). In continuità con tale prospettiva, Martini (2006) evidenzia come la Valutazione venga impiegata per soddisfare cinque principali bisogni conoscitivi: (i) allocare le risorse agli usi considerati più meritevoli o strategicamente rilevanti; (ii) verificare la performance delle organizzazioni e dei servizi, monitorando risultati e livelli di efficienza; (iii) rendere conto delle realizzazioni ottenute in un determinato ambito di azione pubblica; (iv) effettuare un'analisi critica dei processi di attuazione; (v) stimare gli effetti prodotti da una politica, sia quelli intenzionali sia quelli inattesi, positivi o negativi.

richiedendo l'adozione di una postura riflessiva capace di interpretare criticamente i dati e gli output generati dai sistemi di IA, soprattutto quando questi strumenti vengono impiegati nella distribuzione di risorse pubbliche destinate a gruppi sociali vulnerabili e a rischio di esclusione (cfr. par. 4).

Da più parti (Mazzeo Rinaldi, 2017; Bohni Nielsen *et al.*, 2025) si ritiene, infatti, che per comprendere come i processi di automazione stiano riconfigurando le modalità attraverso cui le politiche sociali vengono progettate e implementate sia indispensabile una riflessione che si interroghi sul ruolo che la Valutazione possa assumere nell'orientare le *policies* algoritmiche, facendo propri i seguenti obiettivi: (i) governare in modo consapevole e responsabile gli strumenti di IA; (ii) aprire la *black box* dei modelli socio-tecnici, rendendo trasparenti dati, assunzioni e regole che ne strutturano il funzionamento; (iii) limitare il rischio di amplificare la riproduzione di diseguaglianze radicate e preesistenti.

A partire da queste considerazioni, il saggio si articola in tre sezioni.

La prima sezione (cfr. par. 2) affronta il tema del limitato ricorso all'analisi computazionale nei processi valutativi; una carenza che segnala la necessità di un avanzamento epistemologico e tecnico, indispensabile per orientare l'analisi e l'interpretazione delle evidenze prodotte dai sistemi *data-intensive*.

La seconda sezione (cfr. par. 3) approfondisce le implicazioni metodologiche connesse all'uso dei dati massivi e degli strumenti computazionali, interrogandosi sugli approcci e i metodi valutativi più adeguati a interpretare in modo appropriato le evidenze prodotte dalla *data science*.

La terza sezione (cfr. par. 4) si concentra invece sulla *discriminazione algoritmica*, passando in rassegna alcuni casi emblematici di *bias* sociali prodotti da sistemi socio-tecnici impiegati nell'allocazione di risorse pubbliche. L'analisi di questi esempi mostra come la Valutazione, grazie all'apporto della *teoria dell'azione e della decisione*, sia in grado di aprire la *black box* dei modelli *data-intensive*, rendendo visibili i meccanismi attraverso i quali i sistemi automatizzati possono generare esiti differenziati – e molto spesso discriminatori – sui vari gruppi di interesse.

Questa prospettiva permette di mettere in discussione l'idea di un'"occasione mancata" nel rapporto tra *data science* e Valutazione, mostrando come quest'ultima, per non perdere la capacità di cogliere la complessità dei contesti e degli impatti delle *policies* algoritmiche, debba saper si ibridare con gli strumenti computazionali, acquisendo nuovi linguaggi e misurandosi con nuovi orizzonti analitici.

2. Sistemi *data-intensive* e Valutazione. Note critiche su un rapporto *in fieri*

Muovendo dalla prospettiva degli STS², appare oggi innegabile il ruolo pervasivo del binomio *big data–intelligenza artificiale* nei processi decisionali e nella progettazione di interventi e servizi pubblici (*cf.* par. 4).

Riferendosi in particolare modo alla pratica valutativa, Mazzeo Rinaldi e Bohni Nielsen (2025) sottolineano il significativo ritardo con cui il settore della Valutazione continua a confrontarsi con la richiesta – proveniente dalla *data science* – di integrare l’uso delle tecnologie emergenti nelle diverse fasi del ciclo di vita di una policy (dalla definizione dell’*evaluando* alla formulazione delle domande valutative, fino alla raccolta, archiviazione ed elaborazione dei dati) (Leeuw, 2025; Bruce *et al.*, 2025).

Nel contesto italiano, tale ritardo si accentua e si traduce in un impiego riduttivo e marginale delle scienze computazionali nella Valutazione, molto spesso confinato a sperimentazioni isolate che impediscono al settore di evolvere e di cogliere appieno il potenziale offerto dalla crescita esponenziale dei dati e dal sempre più evidente intreccio tra scienze computazionali e automazione nei processi decisionali (Mazzeo Rinaldi, 2017; Mazzeo Rinaldi e Occhipinti, 2023).

Diversi contributi si sono interrogati sulle ragioni di questo ritardo, individuando due principali ordini di motivazioni.

Un primo fattore riguarda la persistente *mancaza di alfabetizzazione digitale* all’interno della comunità dei valutatori, un gap che ostacola lo sviluppo di un approccio multidisciplinare all’analisi e all’interpretazione dei dati raccolti con le tecnologie emergenti dell’IA.

Petersson e Breul (2017) ritengono che l’avanzamento tecnologico stia trasformando profondamente la Valutazione come campo disciplinare.

Poiché alcuni compiti valutativi vengono automatizzati, altri potenziati e altri ancora completamente reinventati (Eloundou *et al.*, 2023), i valutatori sono chiamati a familiarizzare con gli strumenti digitali e ad acquisire quelle competenze tecniche e informatiche (tipiche della *data science*) che oggi

² La prospettiva degli STS appare rilevante poiché consente di leggere le politiche pubbliche come esiti di *reti socio-tecniche* in cui pratiche, dispositivi e attori (umani e non-umani) risultano intrinsecamente connessi, incidendo sia sui significati delle *policies* sia sull’interpretazione dei loro effetti. In particolare, si ritiene che il lessico analitico proposto da Bruno Latour in *Reassembling the Social* (2005) possa rappresentare uno strumento utile per orientare la Valutazione verso l’adozione di una postura riflessiva, anziché respingente, capace di interrogare criticamente i processi di *blackboxing* che caratterizzano i sistemi *data-intensive*.

risultano centrali per la stessa sopravvivenza della Valutazione come pratica professionale.

Un secondo fattore è riconducibile a forme di *conservatorismo professionale e cognitivo*, legate allo scetticismo verso le tecnologie emergenti basate sui BD e sull'IA, spesso percepite più come una minaccia che come un'opportunità di avanzamento epistemico e metodologico.

In questa prospettiva, se da un lato i BD e le tecnologie emergenti offrono la possibilità di ampliare la comprensione dei contesti entro cui si inscrivono i fenomeni sociali oggetto di osservazione, dall'altro configurano il rischio (ampiamente sconfessato³), della cosiddetta “fine della teoria” (Anderson, 2008), in virtù della progressiva marginalizzazione dei fondamenti teorici che sostengono l'analisi sociale (Kitchin, 2014; Savage e Burrows, 2007).

Inserendosi nel dibattito tra la “promessa” e la “minaccia” insita nei processi di *datafication*, Sabetta (2018) osserva come i BD e la diffusione di strumenti di analisi sempre più sofisticati stiano alimentando la diatriba (ancora irrisolta) tra *integrazione* vs *sostituzione* dei dati rispetto alla teoria.

Nel dibattito sociologico sui BD, questa tensione si traduce, da un lato, nell'entusiasmo per le possibilità esplorative offerte dalle piattaforme digitali (*la promessa*), connesse all'idea che la moltiplicazione delle tracce sul web possa ampliare in modo inedito la comprensione dei fenomeni sociali (Bollier, 2010). Dall'altro lato, emergono posizioni più radicali che considerano i BD una *minaccia* per la riflessività istituzionale della sociologia, ovvero per la sua capacità di ripensare strategie e metodi di ricerca sfruttando appieno le potenzialità degli ambienti digitali. Secondo tali approcci, con un numero sufficiente di dati “i numeri parlano da sé” e la *correlazione* può sostituirsi alla *spiegazione*, fino a prefigurare l'affermazione di un modello di ricerca puramente *data-driven* (Mayer-Schönberger e Cukier, 2013) la cui validità sarebbe garantita unicamente dalla sua capacità predittiva⁴,

³ Il rischio della cosiddetta “fine della teoria”, evocato da Anderson (2008) nelle prime fasi di diffusione dei Big Data, è stato successivamente ampiamente discusso e ridimensionato nella letteratura sociologica. In particolare, è stato mostrato come l'uso intensivo di dati e tecnologie computazionali non comporti necessariamente una marginalizzazione della teoria, ma richieda piuttosto un suo ripensamento e rafforzamento all'interno di nuovi quadri analitici capaci di integrare dati, contesto e interpretazione. In questa prospettiva, i Big Data non sostituiscono la teoria, ma sollecitano approcci teoricamente informati, in grado di interrogare criticamente i processi di produzione dei dati, i presupposti conoscitivi incorporati nelle tecnologie e i meccanismi sociali che sottendono gli esiti osservati.

⁴ La posizione, epistemologicamente fragile, del “basta che funzioni” – espressione steno-grafica coniata da Lorenzo Sabetta (2018) – muove dalla convinzione che i modelli descrittivi basati su raccolta e analisi di BD possano essere considerati strumenti autosufficienti.

senza alcun riferimento all'analisi concettuale e alla logica dell'inferenza (Fasanella, 2024). Un antidoto a questa deriva anti-teorica, che finisce per considerare i modelli esplicativi fondati sull'uso di BD e IA come parte dell'*explanans* (ciò che spiega) anziché dell'*explanandum* (ciò che deve essere spiegato), si coglie nella prospettiva *actor-network* di Latour (2005), per la quale “an actor is what is made to act by many others” (p. 46); i fenomeni sociali, dunque, non preesistono ai dispositivi di osservazione, bensì emergono attraverso assemblaggi di *attori umani* e *non umani* che, interagendo e stabilizzandosi in *reti socio-tecniche*, contribuiscono a co-costruire e co-produrre la realtà osservata. Ne deriva che i BD, lungi dal rivelare verità oggettive, risultano quali strumenti operativi che integrano – senza mai sostituire – le cornici teoriche, le domande di ricerca e la riflessione causale.

Sul piano della Valutazione dei sistemi *data-intensive*, la sfida consiste nel costruire un dialogo epistemico e operativo che riconosca la natura co-costruita dei dati, affiancando il processo euristico di ricostruzione delle teorie implicite con un'analisi attenta sia delle componenti tecniche sia delle dinamiche organizzative e sociali entro cui i sistemi algoritmici vengono progettati e utilizzati. In particolare, una Valutazione orientata alla comprensione critica dei sistemi *data-intensive* deve essere in grado di (i) individuare *norme, valori e meccanismi* incorporati nelle decisioni algoritmiche; (ii) prestare attenzione ai contesti entro i quali tali sistemi prendono forma (dalle relazioni tra sviluppatori, attuatori e valutatori, ai potenziali impatti che le decisioni algoritmiche possono produrre su altre misure concorrenti, fino agli effetti – diretti e indiretti – sui beneficiari) (cfr. par. 4).

La tavola seguente (tav. 1) illustra le principali aree di frizione che caratterizzano la relazione tra Valutazione e tecnologie *data-intensive*.

ti di spiegazione per via delle caratteristiche intrinseche che li contraddistinguono. Secondo Kitchin (2014), la potenza esplicativa dei BD (in specie quelli estratti dalle piattaforme di social media) deriva dai seguenti fattori: l'elevato *volume* delle informazioni raccolte; l'alta *velocità* con cui tali dati vengono generati e trasmessi, spesso in tempo reale; la loro *varietà*, che può comprendere sia dati strutturati sia dati non strutturati; l'*aspirazione alla totalità* delle informazioni disponibili – in termini di estensione; un livello di dettaglio molto fine – in termini di intensione; una marcata *relazionalità strutturale*, che consente di combinare più dataset; infine, *flessibilità e scalabilità*, che permettono ai dati di espandersi rapidamente.

Tav. 1 – Aree di frizione tra Valutazione e sistemi data-intensive

<i>Aree di frizione</i>	<i>Caratteristiche della Valutazione</i>	<i>Caratteristiche dei sistemi data-intensive</i>	<i>Effetti/Esempi di frizione</i>
<i>Differenze epistemologiche e metodologiche</i>	Approcci causali, qualitativi, partecipativi; attenzione al contesto e ai valori	Approcci predittivi, quantitativi, automatizzati	Difficile integrazione delle analisi predittive nei modelli interpretativi delle policy e di lettura dei risultati
<i>Mancaza di linguaggi e strumenti condivisi</i>	Focus sul contesto e sull'empowerment degli attori	Focus su performance, efficienza e accuratezza, dei sistemi di automazione	Incomprensione reciproca tra sviluppatori IA e valutatori
<i>Marginalità della valutazione nel ciclo dell'IA</i>	Presenza tardiva o assente nei processi di design e sviluppo delle tecnologie	Progresso tecnologico rapido e non sottoposto ad accurato monitoraggio e valutazione	Policy IA non monitorate, esiti sociali imprevisi e discriminazione algoritmica
<i>Opacità e mancanza di trasparenza (black box)</i>	Richiesta di accountability, ricostruibilità dei processi	Modelli ML opachi e non interpretabili	Impossibilità di valutare l'equità e il funzionamento interno dell'algoritmo
<i>Frammentarietà del quadro regolamentare e normativo</i>	Mancaza di linee guida specifiche per valutare l'IA; cultura valutativa debole	Quadro normativo in evoluzione	Incertezza operativa, assenza di obblighi valutativi nei progetti IA pubblici

Su un *piano epistemologico*, la Valutazione – analogamente alla ricerca sociale – si fonda su approcci causali, qualitativi e partecipativi, sensibili al contesto e ai valori degli attori coinvolti. I sistemi *data-intensive*, al contrario, privilegiano logiche predittive, quantitative e automatizzate, che possono rendere complessa l'integrazione di analisi basate su correlazioni all'interno di modelli valutativi orientati alla spiegazione causale (Leeuw, 2025). A ciò si aggiunge la *mancaza di linguaggi e strumenti condivisi*, che genera ulteriori attriti: mentre la Valutazione mira all'*empowerment* degli attori e a una comprensione situata dei processi, l'ingegneria algoritmica opera prevalentemente attraverso indicatori di performance, efficienza e accuratezza. Un'ulteriore area di frizione riguarda la *marginalità della Valutazione nei cicli di progettazione dell'IA*, spesso sviluppata senza monitoraggio o verifica sistematica, con il risultato di policy algoritmiche non valutate, effetti inattesi e potenziali discriminazioni (Eubanks, 2018). Anche

l'opacità dei sistemi automatizzati ostacola l'integrazione tra i due campi dal momento che molti modelli di *machine learning* funzionano come "black box" (Guidotti *et al.*, 2018), impedendo di accertare equità, bias e corretto funzionamento degli algoritmi. Infine, la *frammentarietà del quadro regolamentare e normativo* (Mazzeo Rinaldi, 2017), che non prevede obblighi valutativi chiari nei progetti pubblici di IA, può alimentare incertezza operativa e indebolire la capacità degli attori di governare responsabilmente l'adozione delle tecnologie algoritmiche.

La distanza tra Valutazione e sistemi *data-intensive* non dipende, dunque, soltanto da un gap formativo, ma da divergenze strutturali che riguardano dimensioni epistemologiche, linguaggi professionali e assetti istituzionali. In tale quadro, risulta quanto mai rilevante l'auspicio che i valutatori sviluppino consapevolezza più solide rispetto alle opportunità e ai limiti delle tecnologie *data-intensive*. Ciò richiede la promozione di una reale convergenza di saperi e competenze, sostenendo processi di collaborazione strutturata tra discipline e professioni differenti (Mazzeo Rinaldi e Bohni Nielsen, 2025).

3. Valutazione dei sistemi *data-intensive* e teoria della complessità

Per approfondire il contributo che la Valutazione può fornire alla progettazione e all'attuazione di interventi che adottano misure algoritmiche, è utile richiamare la prospettiva analitica della *teoria della complessità* (Rogers, 2008, 2016), che invita a considerare i programmi pubblici come sistemi dinamici, adattivi e profondamente interconnessi ai contesti (valoriali, sociali e normativi) entro i quali vengono realizzati.

Sempre più spesso, infatti, la Valutazione dei programmi pubblici si confronta con sistemi complessi, caratterizzati da una molteplicità di attori, livelli di *governance* articolati, traiettorie temporali non lineari e processi frammentati che si influenzano reciprocamente. In tali ambiti, l'interazione ricorsiva tra decisioni, vincoli organizzativi e dinamiche operative non solo incide sulle modalità di attuazione degli interventi, ma ridefinisce progressivamente anche i loro esiti attesi, influenzando il modo in cui questi stessi vengono interpretati.

La distinzione proposta da Rogers (2016) tra programmi *semplici*, *complicati* e *complessi*⁵ mostra come i *disegni valutativi di ricerca* che accom-

⁵ I *programmi semplici* sono interventi standardizzati, caratterizzati da obiettivi chiari e da procedure di attuazione uniformi. Vengono implementati da un'unica organizzazione, che

pagnano l'attuazione di programmi complessi – caratterizzati da un'elevata dose di incertezza sui loro esiti, come nel caso delle politiche algoritmiche (cfr. par. 4) – non possano essere concepiti come modelli standard e rigidamente *goal-oriented*. Al contrario, essi devono incorporare i fattori di variabilità ed eterogeneità che contraddistinguono i sistemi socio-tecnici contemporanei, adattandosi alle dinamiche evolutive dei contesti in cui questi prendono forma.

Come argomentaremo più avanti, affinché sia possibile supportare efficacemente l'evoluzione dei programmi verso maggiore coerenza, trasparenza ed equità, è importante che la Valutazione (i) conosca in modo approfondito le assunzioni teoriche, i criteri decisionali e i meccanismi di funzionamento che orientano i sistemi algoritmici; (ii) sia “cucita su misura” dei bisogni che emergono nel corso dell'implementazione dei programmi (Rossi *et al.*, 2007), così da cogliere le implicazioni sociali, valoriali, tecnologiche e istituzionali che influenzano la produzione di esiti previsti e imprevisi.

Volendo sintetizzare, è possibile individuare diverse sfide metodologiche che la Valutazione dei sistemi *data-intensive* è chiamata ad affrontare. Alcune di queste sono tradizionali e riguardano: il *disegno di ricerca*, la *raccolta dei dati* e la loro *analisi* (Mazzeo Rinaldi, 2017; Mazzeo Rinaldi e Bohni Nielsen, 2025).

Con riferimento alla *costruzione dei disegni di ricerca*, la disponibilità di grandi basi di dati può sostenere disegni controfattuali (*randomized control trials*) più accurati e maggiormente rispondenti alle domande valutative, poiché consente di incorporare informazioni molto dettagliate sulle caratteristiche dei partecipanti, sulle condizioni di contesto e sul paniere di variabili da rilevare e approfondire nel corso della ricerca⁶. Tale ricchezza

applica passaggi operativi analoghi in contesti simili; il loro funzionamento è in larga parte prevedibile. I *programmi complicati* sono interventi che presentano componenti multiple o che possono produrre risultati solo se coordinati con altri programmi. La loro implementazione coinvolge diverse organizzazioni, collocate in contesti differenti e caratterizzate da modalità operative eterogenee. La difficoltà, in questo caso, deriva dalla necessità di gestire l'interdipendenza tra le parti. I *programmi complessi*, infine, non possono essere standardizzati, poiché si adattano dinamicamente all'emergere dei bisogni e alla comprensione progressiva di ciò che funziona nei contesti reali entro cui vengono realizzati. Sono implementati da una pluralità di organizzazioni, spesso con ruoli non stabiliti a priori, e i loro risultati dipendono fortemente dalle condizioni di partenza, dalle relazioni tra attori e dalla variabilità dei contesti. La complessità deriva dalla natura non lineare, emergente e situata dei meccanismi che li sostengono.

⁶ Nell'ambito della ricerca valutativa e sociale, l'argomentazione controfattuale consiste nella stima di ciò che sarebbe accaduto agli stessi soggetti in assenza dell'intervento sottoposto ad analisi. La disponibilità di basi di dati molto dettagliate contribuisce a rafforzare questo tipo di argomentazione perché consente di descrivere con maggiore accuratezza le

informativa può migliorare l'analisi d'impatto delle politiche pubbliche, ampliando la capacità della Valutazione di distinguere tra influenze dell'intervento e fattori legati al contesto (Mazzeo Rinaldi, 2017; Picciotto, 2020; York e Bamberger, 2020).

Sul piano della *raccolta dei dati*, i sistemi *data-intensive* consentono di arginare l'elevato costo e la lentezza delle indagini condotte su dati primari, permettendo di raccogliere evidenze su quei gruppi marginali, generalmente difficili da raggiungere attraverso piste tradizionali di ricerca.

Infine, l'evoluzione dei modelli di IA amplia il corredo di strumenti e tecniche di *analisi dei dati*, consentendo l'applicazione di modelli: *descrittivi*, utili a comprendere trend passati; *predittivi*, in grado di stimare scenari futuri e *prescrittivi*, capaci di suggerire, sulla base dei dati disponibili, combinazioni ottimali di misure e modalità di intervento (York e Bamberger, 2020; York, 2025).

L'avvento dei BD e dell'IA sta senz'altro prefigurando sfide nuove per la Valutazione. Una delle innovazioni più rilevanti riguarda la possibilità di costruire valutazioni iterative e adattive⁷, capaci di aggiornare la *teoria del cambiamento*⁸ (Weiss, 2007a, 2007b) durante l'implementazione del pro-

caratteristiche individuali dei partecipanti, le condizioni di contesto e le variabili rilevanti che, potenzialmente, influenzano gli esiti. In questo modo diventa possibile costruire gruppi di controllo più comparabili (ad esempio attraverso procedure di stratificazione, matching o randomizzazione informata), riducendo i problemi di selezione e aumentando la plausibilità del controfattuale.

⁷ In questa prospettiva, nota come *developmental evaluation* (Patton, 2011), la Valutazione segue lo sviluppo dei programmi, accompagnandone l'evoluzione e consentendo un processo di apprendimento costante e iterativo tra sviluppatori, attori e valutatori dei sistemi *data-intensive*. Tale configurazione adattiva della Valutazione è altresì capace di attivare meccanismi di *double-loop learning* (Stame, 2014, 2016) tra gli stakeholders, consentendo un'interazione più strutturata, che contribuisce a ridefinire priorità, strategie e modalità operative dei programmi/sistemi alla luce delle evidenze che emergono nei diversi contesti di attuazione. In questo modo, la Valutazione non si configura come strumento di controllo (ex post), ma come dispositivo dialogico e riflessivo (ex ante), capace di accompagnare l'intero ciclo di vita delle *policies* algoritmiche (in itinere), orientandone l'uso verso configurazioni più eque, socialmente legittime e cucite su misura dei bisogni che emergono nei contesti di implementazione.

⁸ Secondo Carol H. Weiss (2007a, 2007b) la Valutazione non deve limitarsi a misurare gli esiti di un intervento, ma deve ricostruire gli assunti espliciti e impliciti – teorici, valoriali, normativi – che guidano programmi, politiche e interventi pubblici (*teoria del cambiamento*). All'interno della *teoria del cambiamento* (intesa come la trama concettuale che collega attività, meccanismi e risultati attesi di un programma) si distinguono: (i) la *teoria del programma*, che riguarda gli assunti causali alla base dell'intervento, con particolare riferimento ai meccanismi necessari per produrre output e outcome desiderati; (ii) la *teoria dell'implementazione*, che ricostruisce la sequenza concreta di passaggi, attività, decisioni e interazioni attraverso cui gli assunti del programma prendono forma nella realtà e si traducono in soluzioni pratiche.

gramma, ottenendo un ciclo continuo di apprendimento (*double loop learning*) che permette di correggere tempestivamente assunti, ipotesi causali e scelte operative.

D'altra parte, il flusso informativo generato dai sistemi *data-intensive* – reso possibile dalla raccolta e dall'aggiornamento continuativo di grandi basi di dati – permette di focalizzare l'attenzione sull'*implementation gap*, consentendo di identificare adattamenti e deviazioni del programma, nella sua messa in opera, e di riportare gli stessi al modello teorico originario, che risulta così aggiornato alla luce degli sviluppi emergenti e delle trasformazioni che si verificano nel corso della sua attuazione (Protta, 1978; Pressman e Wildavsky, 1973; Ham e Hill, 1986; Maynard-Moody e Portillo, 2010; Brodtkin, 2015).

In questo quadro, la Valutazione si configura come un dispositivo dinamico che accompagna l'azione pubblica, integrandosi “con” e orientando “i” flussi informativi prodotti dai sistemi *data-intensive*, affinché risultino realmente aderenti alle esigenze e ai bisogni del contesto.

York e Bamberger (2025) sottolineano come la *data science* stia profondamente trasformando la pratica della ricerca valutativa⁹ svolta nell'ambito dei progetti di sviluppo e di cooperazione internazionale. In questo contesto, la complessità intrinseca dei programmi – data dalla molteplicità degli attori coinvolti, dalla varietà dei contesti socio-istituzionali e dalle interdipendenze tra diversi livelli di *governance* (locali, nazionali e globali) – richiede l'adozione di strumenti in grado di raccogliere, integrare e analizzare volumi di dati molto più ampi, eterogenei e dinamici rispetto al passato (Mazzeo Rinaldi e Bohni Nielsen, 2025).

Intesa come pratica riflessiva, la Valutazione può svolgere un ruolo di verifica empirica e interpretazione critica dei risultati prodotti dagli algoritmi, mettendoli alla prova attraverso fonti plurali e prospettive complementari, soprattutto quando si tratta di comprenderne gli effetti in situazioni reali. L'attenzione al *dato situato* consente così ai valutatori di riconoscere *bias*, limiti e implicazioni etiche delle scelte algoritmiche, scardinando le rappresentazioni semplificate della realtà sociale che, molto spesso, sono alla loro base.

⁹ In particolare, gli autori evidenziano il contributo dei sistemi *data-intensive* alla: raccolta automatizzata dei dati amministrativi; costruzione di modelli causali (*causal modeling*) basati su tecniche di apprendimento automatico e alla possibilità di realizzare analisi in tempo reale, trasformando la Valutazione da pratica prevalentemente *ex-post* a leva di *management* continuo, capace di accompagnare i processi decisionali lungo l'intero ciclo di vita delle politiche.

4. *Bias, black box* e meccanismi sociali: esempi di discriminazione algoritmica

Come discusso nel paragrafo precedente, i sistemi *data-intensive* offrono alla Valutazione opportunità rilevanti per rafforzare le diverse fasi del lavoro interpretativo – dalla costruzione dei disegni di ricerca alla raccolta e all’analisi dei dati – ampliando significativamente la portata e la profondità delle evidenze disponibili. Tali potenzialità, tuttavia, richiedono di essere affiancate da una riflessione altrettanto rigorosa sui limiti e le principali avvertenze metodologiche connesse ai rischi associati al ricorso ai BD e all’IA. La cosiddetta “fallacia dei grandi numeri” induce infatti a ritenere che la quantità garantisca di per sé qualità, mentre *bias* di selezione e distorsioni strutturali possono essere addirittura amplificati da dataset molto estesi, ancor di più se le fonti informative sono state raccolte per scopi differenti da quelli della Valutazione e dai suoi costrutti di interesse (Mazzeo Rinaldi, 2017).

A proposito dei *bias* distorsivi e dei rischi connessi all’impiego di sistemi automatizzati nei processi decisionali, Leeuw (2025) sottolinea come l’attuale tendenza verso la crescente *algoritmizzazione* delle politiche pubbliche (in particolare quelle che incidono sulle vulnerabilità multiple) stia producendo delle conseguenze rilevanti sul piano dell’equità sociale.

Secondo lo studioso, la ragione risiede nel fatto che, sempre più spesso, i sistemi *data-intensive* vengono progettati e adottati senza predisporre adeguati meccanismi di feedback, né prima né dopo la loro implementazione. In assenza di cicli di apprendimento che mettano in connessione sviluppatori, funzionari e cittadini destinatari dell’IA, il rischio è che le distorsioni generate dall’uso degli algoritmi non vengano individuate né corrette in tempo.

In altri casi, l’applicazione di tali sistemi passa rapidamente dalla fase pilota a quella di attuazione su larga scala, con la conseguenza che le amministrazioni coinvolte finiscono per adottare tecnologie di cui non comprendono pienamente il funzionamento e rispetto alle quali non dispongono delle competenze necessarie per valutarne l’affidabilità e i margini di errore (Eubanks, 2018).

Nel contesto di questa riflessione, il ricorso all’impianto teorico della *Realistic Evaluation* (Pawson e Tilley, 1997) risulta particolarmente fruttuoso, oltre che strategico.

Tale approccio offre una cornice efficace per aprire la *black box* dei sistemi *data-intensive*, permettendo di comprendere come i destinatari delle politiche algoritmiche reinterpretino o reagiscano ad esse e di confrontare tali configurazioni con quelle osservabili in contesti analoghi in cui l’IA non è prevista – o è utilizzata in modo differente. Attraverso la classica

formula *Context+Mechanism=Outcome* (CMO), la *Realistic Evaluation* permette di isolare con maggiore precisione i meccanismi relazionali (e distorsivi) sottesi ai processi di implementazione dei sistemi algoritmici. Su un piano prettamente metodologico, inoltre, fornisce un quadro analitico in grado di combinare (i) l'auditing di dati e modelli, (ii) l'osservazione etnografica *in situ*, (iii) disegni sperimentali e osservazionali, integrando coeentemente approcci quantitativi e qualitativi di ricerca¹⁰.

La meta-analisi condotta da Leeuw (2025) su un corpus selezionato di articoli e documenti dedicati alla valutazione dei sistemi algoritmici permette di mettere meglio a fuoco la natura polisemica e articolata dei meccanismi che tali sistemi attivano una volta adottati. Tali meccanismi sono ricondotti dall'autore a una serie di dimensioni concettuali – *reasoning; social interactions; antropomorphism; emotion handling; attitudinal and perceptual resources* – che rivelano il modo in cui l'artefatto tecnologico orienta aspettative, proiezioni e schemi d'azione.

La dimensione del ragionamento (*reasoning*) riguarda i processi cognitivi e le condizioni materiali che orientano l'adozione e l'uso dell'IA (tra queste: la compatibilità dell'innovazione con le pratiche esistenti e la percezione di fairness dei modelli). Nella dimensione interazionale (*social interactions*) è decisiva la qualità del coordinamento tra persone, ruoli e si-

¹⁰ Un esempio efficace dell'approccio realistico, ripreso da Leeuw (2025), è offerto da una Valutazione sul ruolo dell'IA utilizzata in un protocollo di chirurgia robot-assistita. Lo studio mostra come la *Realistic Evaluation* consenta di ricostruire in modo situato i processi di implementazione del protocollo medico, combinando metodi qualitativi e quantitativi in un'unica struttura analitica. La Valutazione prende avvio: i dalla ricostruzione della *teoria del programma* alla base del protocollo chirurgico, attraverso un lavoro preliminare di interviste in profondità al team di sala operatoria, volte a identificare i meccanismi attesi dall'introduzione dell'IA (ottimizzazione dei tempi, risultati nel breve-medio periodo). In seguito, si passa (ii) alla ricostruzione della *teoria dell'implementazione*, realizzata attraverso osservazioni etnografiche multi-sito. L'obiettivo è verificare le diverse modalità con cui il protocollo medico viene effettivamente applicato nei vari contesti ospedalieri, mettendo a fuoco come i meccanismi socio-relazionali si attivino (o falliscano) in relazione alle condizioni specifiche di ciascun setting organizzativo (dotazioni tecnologiche, cultura professionale, routine operative, dinamiche di coordinamento tra équipe). In una fase successiva, l'indagine viene estesa ad altre discipline chirurgiche, per valutare la trasferibilità del protocollo e comprendere se e come i meccanismi attivati si manifestino in specialità differenti. Il percorso si conclude con un workshop con gli stakeholder, in cui vengono discussi risultati, interpretazioni, criticità emerse e implicazioni pratiche, così da integrare prospettive professionali diverse e validare il modello finale. Il caso di studio mostra con chiarezza come l'approccio *realista* consenta di aprire la *black box* delle tecnologie algoritmiche, identificando i meccanismi generati dall'IA, mettendoli in relazione con i contesti specifici in cui si attivano e promuovendo un dialogo aperto e inter-stakeholder tra sviluppatori, attori, valutatori e beneficiari delle tecnologie.

stemi automatizzati. Le diverse configurazioni *human-in/on/off-the-loop* possono, talvolta, ridurre il peso decisionale dei destinatari – diretti e indiretti – a cui tali sistemi si rivolgono. In altre circostanze, invece, configurazioni più mature ed efficaci di cooperazione uomo-macchina sostengono relazioni maggiormente fluide ed efficaci. Nella dimensione dell’antropomorfismo (*antropomorphism*) rientrano i processi di attribuzione di caratteristiche umane a sistemi di IA, considerati entità dotate di intenzioni, emozioni o capacità relazionali. Questa proiezione può facilitare lo sviluppo di forme di vicinanza emotiva e di relazioni parasociali con la tecnologia, influenzando significativamente i comportamenti degli utenti e il modo in cui questi interpretano le decisioni algoritmiche. Nella dimensione delle emozioni e dell’empatia (*emotion handling*) sono ricomprese le forme di auto-consapevolezza emotiva, i giudizi e i comportamenti che gli utenti sviluppano nell’interazione con i sistemi di IA¹¹. Infine, la dimensione delle risorse attitudinali e percettive (*attitudinal and perceptual resources*) ricostruisce i meccanismi che spiegano come atteggiamenti, credenze e percezioni degli utenti influenzino l’accettazione o il rifiuto dei sistemi di IA.

Il ricorso a sistemi automatizzati nell’erogazione di misure di assistenza rivolte a gruppi sociali vulnerabili può attivare, in determinate circostanze, alcuni dei meccanismi individuati da Leeuw. Ciò avviene, ad esempio, quando l’antropomorfizzazione delle tecnologie digitali alimenta un’eccessiva fiducia nelle decisioni assunte da infrastrutture digitali avanzate, riducendo la propensione ad esercitare un controllo critico sul funzionamento di tali sistemi, e favorendo forme di implementazione meccanica delle prescrizioni algoritmiche.

Tra i meccanismi potenzialmente legati all’interazione uomo-macchina è utile includere il *moral hazard*¹², da interpretare qui come la tendenza degli attuatori a “profittare” delle asimmetrie informative prodotte dall’adozione di sistemi decisionali opachi, che funzionano come *black box*. Ri-

¹¹ Le risposte emotive prodotte dall’interazione uomo-macchina plasmano profondamente fiducia, disponibilità all’uso e percezione di equità delle decisioni automatizzate.

¹² Nel quadro dell’analisi organizzativa (Crozier e Friedberg, 1977), il *moral hazard* designa la situazione in cui un attore modifica il proprio comportamento perché sa di poter contare su una qualche forma di copertura o protezione rispetto alle conseguenze delle proprie azioni. Nelle organizzazioni complesse, tale dinamica può essere amplificata da procedure standardizzate o da automatismi che fungono da scudo tecnico, attenuando la responsabilità individuale e rendendo meno trasparente la catena causale tra azione e risultato. In presenza di sistemi algoritmici – già caratterizzati da opacità strutturale – il *moral hazard* può tradursi in un ulteriore spostamento dell’*accountability* verso la tecnologia, con rischi di deresponsabilizzazione dei decisori umani e consolidamento di pratiche che sfuggono ad una valutazione critica dei dati e delle decisioni.

prendendo gli assunti centrali della *teoria dell'azione e della decisione* (razionalità limitata, *slack* organizzativo, sistemi *loosely coupled* – Simon, 1947; trad.it. 2001; March e Olsen, 1989; trad. it. 1992; Crozier e Friedberg, 1977¹³), risulta evidente quanto l'opacità algoritmica possa rendere più difficile tracciare il nesso tra input, processi e output, ampliando o riducendo gli spazi di discrezionalità decisionale degli attuatori a seconda dei contesti organizzativi e delle *routines* operative in cui tali sistemi vengono incorporati. Considerare i sistemi algoritmici alla stregua di attori opachi, ancorché inseriti in ecologie decisionali complesse e multi-stratificate (per via della pluralità dei livelli decisionali, dagli sviluppatori, agli implementatori agli utenti finali), permette, dunque, di evidenziare quali vincoli e logiche di azione emergono dall'interazione tra architetture tecniche e contesti organizzativi, inclusi gli effetti non intenzionali e le forme di discriminazione che possono derivarne.

Seguendo Virginia Eubanks (2018), la *discriminazione algoritmica* può essere interpretata come l'esito dell'incorporazione automatizzata di pattern di sorveglianza e classificazione che orientano i dispositivi normativo-digitali¹⁴, trasformandoli in ambienti “a basso tasso di diritti”, ovvero in spazi in cui la trasparenza è limitata, la possibilità di ricorso fortemente ridotta e la vulnerabilità degli utenti ulteriormente amplificata.

¹³ L'analisi delle decisioni automatizzate può beneficiare dell'apporto dei teorici dell'azione e della decisione, citati nel testo. In particolare, la categoria analitica della *razionalità limitata* mostra come i decisori – umani e non umani – si accontentino spesso di soluzioni “sufficientemente razionali” anziché pienamente ottimali (March e Olsen, 1989; trad. it. 1992). Questo comporta l'adozione di standard operativi che, lungi dall'essere neutri, possono cristallizzare routine e produrre asimmetrie informative. A ciò si collega il concetto di *slack organizzativo*, inteso come riserva tecnica e valoriale che può, in alcuni casi, attenuare il conflitto e favorire apprendimento e, in altri casi, irrigidire l'organizzazione attorno a criteri definiti *ex-ante* e poco negoziabili (Crozier e Friedberg, 1997). Infine, la prospettiva dei sistemi *loosely coupled* mette in luce l'indebolimento del nesso lineare tra decisione e attuazione, tipico delle architetture algoritmiche, in cui la ridotta trasparenza delle logiche decisionali può amplificare i rischi di discrepanze esecutive nonché favorire la riproduzione di *bias* e pregiudizi che coinvolgono gruppi sociali svantaggiati e a rischio di marginalità (Dentale, 2022).

¹⁴ Per restare nell'ambito delle riflessioni condotte sulla discriminazione algoritmica nei sistemi di *welfare*, è particolarmente interessante il rapporto di ricerca curato da Kevin De Liban (2024), fondatore di *TechTonic Justice* – Organizzazione Non Profit nata per sostenere, insieme alle persone a basso reddito danneggiate dall'IA – battaglie di giustizia e tutela dei diritti. Il rapporto mostra come l'IA sia diventata, ormai, un'infrastruttura centrale nei processi decisionali che regolano la vita delle persone a basso reddito, producendo forme pervasive di discriminazione sistemica. Il rapporto mostra infatti come i sistemi *data-intensive* abilitino funzioni potenzialmente discriminatorie (automatizzare controlli invasivi, imporre carichi amministrativi ingestibili alle famiglie), sul piano sociale, etico e politico.

Nel ricostruire i meccanismi sociali connessi al fenomeno in oggetto Eubanks (2018) mostra come due misure di politica pubblica (statunitense) – il programma federale *Temporary Assistance for Needy Families* (TANF) e l'*Allegheny Family Screening Tool* (AFST) – traducano la storica questione della *deservingness* in architetture digitali che standardizzano la logica della sorveglianza e della penalizzazione.

Nel caso del TANF, intervento federale di sostegno al reddito rivolto alle famiglie in condizione di povertà, la razionalizzazione amministrativa introdotta tramite procedure informatiche (come le sanzioni comminate in via automatica per la “mancata cooperazione” delle famiglie prese in carico alle attività di inserimento sociale) ha contribuito a trasformare le inadempienze¹⁵ degli assistiti in tagli immediati delle prestazioni economiche, senza accertarne le cause o considerare la specificità delle situazioni di disagio.

Questo processo ha di fatto minimizzato il margine di valutazione discrezionale¹⁶ degli operatori, tradizionalmente indispensabile per contestualizzare le situazioni di povertà, rendendo la decisione amministrativa (relativa all’esclusione dal beneficio economico) opaca, difficilmente contestabile e sostanzialmente automatizzata.

Un esito affine si osserva nel caso dell’AFST, un sistema predittivo adottato dal *Department of Human Services* della Contea di Allegheny (Pittsburgh, Pennsylvania) per supportare la fase preliminare di Valutazione delle segnalazioni di sospetto maltrattamento minorile. Il modello elabora un ampio insieme di informazioni amministrative archiviate nel *data warehouse* della contea – tra cui precedenti interazioni con i servizi sociali, dati relativi ad arresti, accesso a servizi sanitari o scolastici, utilizzo di misure di assistenza economica e perfino elementi giudiziari riferibili a più generazioni familiari – e li sintetizza in un punteggio di rischio utilizzato per decidere se avviare o meno un’indagine. Pur presentato come uno strumento

¹⁵ Il mancato svolgimento di un colloquio obbligatorio con i servizi sociali, previsto ai fini della presa in carico o dell’inserimento lavorativo, poteva comportare la decadenza immediata dal beneficio economico.

¹⁶ Questo aspetto è stato rilevato anche da Dentale (2022) nell’analisi dedicata alla valutazione dell’agire discrezionale degli operatori sociali impegnati nell’erogazione di diverse misure di reddito minimo in Italia, con un focus sul caso romano. Lo studio ricostruisce diversi stili di esercizio della discrezionalità decisionale, soffermandosi in particolar modo sullo *stile* rigidamente *procedimentale*, caratterizzato dalla tendenza degli operatori sociali a seguire in modo meccanico le procedure amministrative, riducendo lo spazio per considerare le specificità delle situazioni familiari e individuali. Sebbene non siano riferiti a contesti caratterizzati da procedure automatizzate, questi risultati mettono in luce come l’eccesso di formalismo procedurale possa ostacolare la capacità degli operatori di valutare in modo situato e dialogico le condizioni di vita dei beneficiari inclusi in misure di contrasto alla povertà.

capace di garantire maggiore coerenza e oggettività nelle decisioni, l'AFST si basa su dati raccolti in contesti segnati da pratiche storiche di sovra-sorveglianza delle comunità povere e razzializzate. Ciò comporta che tale sorveglianza differenziale venga tradotta in indizi di rischio apparentemente neutri, trasformando la vulnerabilità in una sorta di attributo ereditario incorporato nei dati. In questo modo l'algoritmo tende a riprodurre e amplificare le asimmetrie già presenti nelle istituzioni, orientando l'intervento pubblico attraverso una "matematizzazione del sospetto" che riflette più la storia dei controlli esercitati sulle famiglie che i loro bisogni effettivi.

Letti attraverso la lente della *teoria dell'azione e della decisione* (cfr. *supra*), questi due casi mostrano con forza che la *discriminazione algoritmica* non ha una natura neutrale né esclusivamente tecnica, ma deriva dalle scelte politiche, morali e culturali incorporate in sistemi digitali che, molto spesso, fungono da *black box*.

Nel TANF, l'automatizzazione delle procedure e la compressione dei margini discrezionali dei funzionari configurano un meccanismo di *means-ends decoupling* (Jabbouri *et al.*, 2019), in cui la tecnologia non funge da supporto all'interazione tra beneficiari e operatori, ma ne diventa il sostituto, trasformandosi in un mero strumento procedurale che allontana l'azione dal fine (sostenere, non già punire, le famiglie in difficoltà). Tale forma di scollamento risulta particolarmente probabile in sistemi altamente frammentati, dove gli attuatori delle decisioni automatizzate operano (a volte consapevolmente) in ambienti opachi, segnati da *causal complexity*, *behavioral invisibility* e *practice multiplicity* (Wijen, 2014). In tali contesti, è plausibile ipotizzare che gli operatori incaricati di applicare le procedure algoritmiche continuino ad attenersi ai protocolli anche quando ne riconoscono l'impatto limitato – o potenzialmente distorsivo – sugli esiti (Bromley e Powell, 2012).

Nel caso dell'AFST, la relazione tra operatori e sistemi algoritmici assume la forma di una cooperazione sostanzialmente asimmetrica: l'algoritmo fornisce un punteggio che orienta l'interpretazione professionale e incide sulle pratiche decisionali persino quando gli operatori mantengono un ruolo formale di supervisione (*human-in-the-loop*). Non si tratta della sostituzione delle categorie cognitive del decisore umano, bensì dell'etero-direzione delle sue valutazioni mediante un meccanismo di *sovra-sorveglianza* (incorporata nei dati di screening) che agisce a danno delle comunità razzializzate¹⁷.

¹⁷ Questo meccanismo viene riportato anche nello studio di Corbett Davies *et al.* (2017), dove si mostra come gli strumenti di *risk assessment* nelle decisioni giudiziarie preprocessuali possano riprodurre o amplificare disparità razziali, classificando con maggiore frequenza come "ad alto rischio" gli imputati neri e compromettendo così l'equità del processo decisionale.

5. Per concludere: *matters of fact or matters of concern?*

La riflessione fin qui condotta mette chiaramente in evidenza come la Valutazione risulti profondamente rinnovata dall'incontro con la *data science* e i sistemi *data-intensive* (cfr. par. 3).

L'analisi dei diversi casi di discriminazione algoritmica ha mostrato che un uso a-critico di questi sistemi – non accompagnato da adeguate procedure di monitoraggio e Valutazione – possa contribuire ad attivare meccanismi di esclusione sociale tanto più pervasivi quanto più opachi risultano i modelli tecnico-matematici posti alla loro base. Affiancando la progettazione delle *architetture algoritmiche*, la Valutazione può contribuire, dunque, a rendere visibili i meccanismi sociali attivati dai modelli *data-intensive*, intercettando tempestivamente gli effetti inattesi e ricomponendo in un quadro unitario tutti quei fattori latenti – valori, opinioni e teorie implicite – che altrimenti rimarrebbero confinati nella *black box* sociotecnica dei programmi.

In questo scenario, l'attenzione al contesto – secondo una matrice *realista* (Pawson e Tilley, 1997) – può contribuire a orientare lo sguardo analitico verso l'elaborazione e l'interpretazione di quei dati che, proprio in quanto situati e iscritti in reti socio-tecniche complesse, richiedono di essere letti nella loro relazione con vincoli di natura politica e sociale.

Si tratta di un approccio che si colloca ben lontano dai rischi della *end of theory* e della “*fine della causalità*” denunciati da alcuni detrattori dei BD e dell'IA, poiché riafferma la necessità di comprendere come i fenomeni sociali oggetto di interesse si inseriscano e si co-producano all'interno di configurazioni miste, abitate da attori umani e non umani. In questa prospettiva, l'investimento nella cultura della Valutazione deve sostenere, da un lato, lo sviluppo di nuovi approcci analitici fondati sui linguaggi computazionali e, dall'altro, misurarsi con le sfide epistemologiche già poste da Charles Wright Mills nel 1959 e riprese dai teorici della società digitale (Lupton, 2018).

Nella sua opera monumentale, *The Sociological Imagination*, Mills denunciava il rischio che una ricerca sociale ridotta alla mera applicazione di modelli tecnici potesse scivolare verso un “*empirismo astratto*” incapace di interrogare criticamente la realtà e incline a confondere strumenti e oggetti di studio. Queste preoccupazioni risultano straordinariamente attuali oggi che la diffusione di tecniche computazionali tende ad appiattire la ricerca sociale sulle due sole fasi dell'*analisi dei dati* (vista la disponibilità massiva di BD) e della *comunicazione dei risultati*, con il rischio che modelli causazionali, perfettamente congegnati dal punto di vista tecnico, continuino a “girare” pur con dosi omeostatiche di teoria (Campelli, 2004).

In questo quadro, il richiamo all'*actor-network theory* di Bruno Latour (2005) contribuisce a rafforzare la validità della riflessione critica già evocata da Mills, spostando l'attenzione dalla concezione dei BD come strutture neutre alle relazioni socio-tecniche attraverso cui essi vengono prodotti, mobilitati e interpretati. L'invito latouriano a superare un'idea sostanzialista del sociale, concentrandosi sui processi di associazione tra elementi eterogenei – bisogni, aspettative proiettate nei dispositivi tecnici e modelli interpretativi – mostra, infatti, come *dati* e *significati* emergano dall'intreccio dinamico di tali relazioni. In tal senso, la distinzione tra *matters of fact* e *matters of concern* (Latour, 2005) evidenzia come la ricerca sociale non debba limitarsi alla registrazione di fatti, bensì provare ad elaborare una cornice teorica che sia capace di individuare le questioni che mobilitano gli *attori sociali* e *non*, cogliendo il significato (sociale, politico ed etico) delle pratiche che ne derivano.

Intorno a tale snodo, la Valutazione dei sistemi *data-intensive*, intesi come assemblaggi socio-tecnic (di valori e pratiche che emergono dall'intersezione di infrastrutture umane e non umane), può offrire un contributo decisivo, permettendo di ricostruire *come* tali sistemi si costituiscano, quali meccanismi attivino e quali effetti producano nei vari contesti in cui operano.

Riferimenti bibliografici

- Amaturo E., Aragona B. (2019), “Per un’epistemologia del digitale: note sull’uso di big data e computazione nella ricerca sociale”, *Quaderni di Sociologia* 81(3), 67-82.
- Anderson C. (2008), *The End of Theory: The Data Deluge makes the Scientific Method Obsolete*, *Wired*, June 23.
- Bohni Nielsen S., Mazzeo Rinaldi F., Petersson G.J. (Eds.) (2025), *Artificial Intelligence and Evaluation: Emerging Technologies and their Implications for Evaluation*. Routledge, New York and London.
- Bollier D. (2010), *The Promise and Peril of Big Data*, The Aspen Institute Communications and Society Program, Washington D.C., The Aspen Institute.
- Brodkin E.Z. (2015), “The Inside Story: Street-level Research in the UD and Beyond”, in Hupe P., Hill M., Buffat A. (Eds.), *Understanding Street-level bureaucracy*, Policy Press, University of Bristol.
- Bromley P., Powell W. (2012), “From Smoke and Mirrors to Walking the Talk: Decoupling in the Contemporary World”, *Academy of Management Annals*, 6, 483–500.
- Bruce K., Gandhi V.J., Vandelanotte J. (2025), “Emerging Technology and Evaluation in International Development”, in Bohni Nielsen S., Mazzeo Rinaldi F., Petersson G.J. (Eds.), *Artificial intelligence and evaluation: Emerging technologies and their implications for evaluation*. Routledge, New York and London.

- Campelli E. (2004), “Per una spiegazione di medio raggio. Ancora sul problema della spiegazione in sociologia”, *Sociologia e ricerca sociale*, XXV, 75, 113-155.
- Crozier M., Friedberg E. (1977), *L'Acteur et le Système*, Seuil, Paris.
- Davies S. C., Pierson E., Feller A., Goel S., Huq A. (2017), “Algorithmic Decision Making and the Cost of Fairness”, *Proceedings of KDD '17*, 10 pp. DOI: 10.1145/3097983.3098095.
- De Liban K. (2024), *Inescapable AI. The Ways AI Decides How Low-Income People Work, Live, Learn, and Survive*, Technotonic Justice.
- Dentale M. (2022), *Condizionalità e discrezionalità nelle politiche sociali. Una ricerca sulla valutazione delle misure di contrasto alla povertà*, FrancoAngeli, Milano.
- Eloundou T., Manning S., Mishkin P., Rock D. (2023), “GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models”, *Paper arXiv*: 2303.10130.
- Eubanks V. (2018), *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*, New York: Picador, St Martin's Press.
- Fasanella A. (2024), “Teoria, concetti, variabili”, in Fasanella A., Mauceri S., Nobile S. (a cura di), *Metodologia della ricerca sociale. Approcci, strategie e tecniche di indagine*, FrancoAngeli, Milano.
- Guidotti R., Monreale A., Ruggieri S., Turini F., Pedreschi D., Giannotti F. (2018), “A Survey of Methods for Explaining Black Box Models”, *ACM Computing Surveys*, 51(5), 1–45. <https://doi.org/10.1145/3236009>
- Jasanoff S. (2012), Genealogies of STS. *Social Studies of Science*, 42(3), 435-441.
- Ham C., Hill M. (1986), *Introduzione all'analisi delle politiche pubbliche*, il Mulino, Bologna.
- Jabbouri R., Truong Y., Schneckenberg D., Palmer M. (2019), *Organizational Decoupling: A Systematic Literature Review and Directions for Future Research*. In EURAM Conference: EURAM 19 Lisbon Portugal 26/06/2019-28/06/2019 (EURAM Proceedings).
- Kitchin R. (2014), “Big Data, New Epistemologies and Paradigm Shifts”, *Big Data & Society*, 1, 1, 1-12.
- Mayer-Schönberger V., Cukier K. (2013), *Big Data: A Revolution that Will Transform how we Work, Live, and Think*, New York, Houghton.
- Maynard-Moody S., Portillo S. (2010), “Street-level Bureaucracy Theory”, in Durant R.F. (Eds.), *Oxford Handbook of American Bureaucracy*, Oxford University Press, Oxford, UK.
- March J.G., Olsen J.P. (1989), *Rediscovering Institutions. The Organizational Basis of Politics*, The Free Press, New York (trad. it.: *Riscoprire le Istituzioni. Le basi organizzative della politica*, il Mulino, Bologna, 1992).
- Mills C.W. (1959), *The Sociological Imagination*, Oxford University Press, Oxford.
- Latour B. (2005), *Reassembling the Social. An Introduction to Actor-Network-Theory*, Oxford University Press, New York.
- Leeuw F.L. (2025), “The Algorithmization of Policy and Society. The Need for a

- Realist Evaluation Approach”, in Bohni Nielsen S., Mazzeo Rinaldi F., Petersson G.J. (Eds.), *Artificial Intelligence and Evaluation: Emerging Technologies and their Implications for Evaluation*. Routledge, New York and London.
- Lupton D. (2018), *Sociologia digitale*, Pearson (a cura di Santoro M. e Timeto F.).
- Martini A. (2006), “Cinque modi di interpretare la valutazione”, in Martini A., Mo Costabella L., Sisti M. (a cura di), *Valutare gli effetti delle politiche pubbliche: metodi e applicazioni al caso italiano*, Formez.
- Mazzeo Rinaldi F. (2017), “Big Data e Valutazione: una relazione ancora da costruire”, *Rassegna italiana di valutazione*, XXI, 68. DOI: 10.3280/RIV2017-068002.
- Mazzeo Rinaldi F., Occhipinti O. (2023), “Big Data, Intelligenza Artificiale e Valutazione: cosa accade in Italia”, *Rassegna italiana di valutazione*, XXVII, 85-86. Doi 10.3280/RIV2023-0850010
- Mazzeo Rinaldi F., Bohni Nielsen S. (2025), “Artificial Intelligence. Challenges for Evaluators”, in Bohni Nielsen S., Mazzeo Rinaldi F., Petersson G.J. (Eds.), *Artificial Intelligence and Evaluation: Emerging Technologies and their Implications for Evaluation*, Routledge, New York and London.
- Palumbo M. (2001), *Il processo di valutazione. Decidere, programmare, valutare*, FrancoAngeli, Milano.
- Patton M. (2011), *Developmental Evaluation*, The Guilford Press, New York.
- Pawson R., Tilley N. (1997), *Realistic Evaluation*, Sage, London.
- Petersson G.K., Breul J.D. (Eds.) (2017), “Cyber Society, Big Data and Evaluation”, *Comparative Policy Evaluation* (Vol. 24). Transaction, New Brunswick, NJ.
- Picciotto R. (2020), “Evaluation and the Big Data Challenge”, *American Journal of Evaluation*, 41(2), 166-181.
- Pressman J., Wildawsky A. (1973), *Implementation*, University of California Press, Berkeley.
- Prottas J.M. (1978), “The Power of the Street-Level Bureaucrat in Public Service Bureaucracies”, *Urban Affairs Quarterly*, 13/3, 285-312.
- Rogers, P. (2008), “Using Programme Theory for Complicated and Complex Programmes”, *Evaluation*, 14(1), 29-48.
- Rogers P. (2016), “Understanding and Supporting Equity: Implications of Methodological and Procedural Choices in Equity-focused Evaluations”, in Donaldson S., Picciotto R. (Eds.), *Evaluation for an Equitable Society*, Information Age Publishing, Charlotte (NC).
- Rossi P.H, Freeman H.E., Lipsey M.W. (2007), “Costruire le valutazioni su misura”, in Stame N. (a cura di), *Classici della Valutazione*, FrancoAngeli, Milano.
- Sabetta L. (2018), “Basta che funzioni? Note su Big Data e teoria sociologica”, *Sociologia Italiana*, 11, 187-194.
- Savage M., Burrows R. (2007), “The Coming Crisis of Empirical Sociology”, *Sociology*, XLI, 5, 885-889.
- Savage M., Burrows R. (2009), “Some Further Reflections on the Coming Crisis of Empirical Sociology”, *Sociology*, XLIII, 4, 762-772.

- Simon H. A. (1947), *Administrative Behavior*, Macmillan, New York (trad. it.: *Il comportamento amministrativo*, il Mulino, Bologna, 2001).
- Stame N. (a cura di) (2007), *Classici della valutazione*, FrancoAngeli, Milano.
- Stame N. (2010), “What doesn’t Work? Three Failures, many Answers”, *Evaluation*, 16(4), 371-387.
- Stame N. (2014), “Positive Thinking Approaches to Evaluation and Program Perspectives”, *Canadian Journal of Program Evaluation*, XXIX, 2, 67-86. doi: 10.3138/cjpe.29.2.67.
- Stame N. (2016), *Valutazione pluralista*, FrancoAngeli, Milano.
- York P., Bamberger M. (2020), *Measuring Results and Impact in the Age of Big Data*, Rockefeller Foundation, New York.
- York P. (2025), “The Future of Evaluation Analytics. Case Studies of Structural Causal Modeling in Action”, in Bohni Nielsen S., Mazzeo Rinaldi F., Petersson G. J. (Eds.), *Artificial intelligence and evaluation: Emerging technologies and their implications for evaluation*, Routledge, New York and London.
- York P., Bamberger M. (2025), “The Applications of Big Data to Strengthen Evaluation”, in Bohni Nielsen S., Mazzeo Rinaldi F., Petersson G.J. (Eds.), *Artificial Intelligence and Evaluation: Emerging Technologies and their Implications for Evaluation*. Routledge, New York and London.
- Weiss C.H. (2007a), “La valutazione basata sulla teoria. Passato, presente e futuro”, in Stame N. (a cura di), *Classici della Valutazione*, FrancoAngeli, Milano.
- Weiss C.H. (2007b), “Abbiamo imparato nulla di nuovo sull’uso della valutazione?”, in Stame N. (a cura di), *Classici della Valutazione*, FrancoAngeli, Milano.
- Wijen F.H. (2014), “Means versus Ends in Opaque Institutional Fields: Trading off Compliance and Achievement in Sustainability Standard Adoption”, *Academy of Management Review*, 39, 302-323.

5. *Il data journalism tra pretesa di “verità” e controversia politico-culturale*

di *Rolando Marini, Matteo Gerli**

1. **Il rapporto tra scienza, dati e giornalismo: una ricognizione**

Il presente contributo intende proporre una riflessione sul rapporto tra uso dei dati e giornalismo, cercando di collegare le implicazioni di natura epistemologica con le trasformazioni del lavoro giornalistico, con il sistema delle interdipendenze in cui questo è coinvolto e con i processi di cambiamento tecnologico che investono la professione.

Il legame tra mezzi di comunicazione (definiti “di massa”) e scienza appare come tratto caratteristico della modernità (McLuhan, 1962; Thompson, 1995), non solo attraverso la progressiva democratizzazione della conoscenza che si verifica con il libro a stampa, proprio nella fase dell’affermazione della scienza moderna, ma anche, successivamente, con la sua divulgazione a opera delle riviste scientifico-letterarie e dei quotidiani. Mentre le prime sono strettamente collegate alle comunità scientifiche e alle loro istituzioni (circoli culturali, accademie e società scientifiche, oltre alle università), offrendo a queste una rete infrastrutturale d’interscambio e adesione allo spirito innovativo (Habermas, 1962; Marino, 1974), i secondi ne assorbono il codice genetico all’interno delle proprie logiche di rappresentazione della realtà.

Più esattamente, il giornalismo – già dai primi decenni del secolo scorso – costruisce uno specifico ideale normativo, incentrato sulla condivisione del metodo scientifico di origine positivista (Lippmann, 1922). L’ideale dell’obiettività, infatti, diventa precetto deontologico e soprattutto operativo di una professione che si istituzionalizza con l’impegno e la pretesa di svincolarsi, nello stesso tempo, dall’ideologia politica, dalle pressioni di attori

* Il capitolo è il risultato della riflessione congiunta degli autori.

esterni (politici, militari ed economici) e dal soggettivismo letterario (Siebert *et al.*, 1956; Schudson, 1995; Hartsock, 2000; Mancini, 2002; Sapiro, 2003).

Ciò si afferma, specialmente nel mondo anglosassone, come tendenza a far assumere al giornalismo la postura delle scienze, allontanandolo dalla presunta e temuta discrezionalità interpretativa, in qualsiasi modo prodotta. Entro tale prospettiva, l'ancoraggio della produzione informativa a dati statistici sembra quindi avere spostato la stessa legittimazione della funzione giornalistica di rappresentazione della realtà e di controllo del potere verso una fonte ritenuta impersonale: fonte che una sedimentata ideologia della professione e il senso comune dei destinatari concepiscono come politicamente neutra rispetto alle possibili contese, capace di irradiare verso la società il suo arbitrato sulla "verità" (Marini, 2014). Si tratta in sostanza di un esito tipico della dominanza del pensiero scientifico e con esso del «linguaggio matematico» nell'evoluzione della cultura nella società contemporanea (Beretta, 2017).

L'uso dei dati nell'informazione diviene una delle *forme di oggettivazione* del resoconto e della lettura della realtà, che giustifica e rende accettabili i processi d'interpretazione, considerati imparziali poiché scientificamente fondati. Ciò si basa, da un lato, sulla visione che fa coincidere la conoscenza scientifica con la verità, e, dall'altro, conseguentemente, sulla necessità di coinvolgere in questo dovere di verità tutte le professioni implicate nella produzione di conoscenza, incluso il giornalismo, chiamato ad assimilarsi alle professioni scientifiche. Si basa – aggiungiamo – sul radicato e spesso risorgente equivoco tra certezza e verità (Watkins, 1986), ovvero sulla inadeguata e fuorviante comprensione delle possibilità di certezza offerte dal procedere della scienza e sulla simmetrica richiesta di verità univoca rivolta alle forme della produzione di notizie (che solo successivamente verranno chiamate *forme di rappresentazione* della realtà).

L'idea dell'oggettività dei dati, insomma, si fonda sulla «matematizzazione» propria del pensiero scientifico contemporaneo (Beretta, 2017), da cui si genera la mancanza di consapevolezza (risalente alla fase positivista delle scienze sociali) della natura del dato come costruito (Marradi e Gasperoni, 2002; Marradi, 2007), ossia come risultato – oltre che di tecniche di rilevazione ed elaborazione – di strategie e disegni di ricerca orientati da finalità negoziate, direttamente e indirettamente, da una pluralità di attori, e non esenti dal gioco spesso asimmetrico delle influenze sociali. I processi di costruzione sociale dei dati non possono renderli oggettivi e corrispondenti a un'idea di verità. Poi, piuttosto che univocamente determinabili, le interpretazioni dei dati sono aperte e suscettibili di discussione.

Non va inoltre dimenticata la presenza nello spazio pubblico di dati che

sono esito di processi di produzione ed elaborazione amministrativa, come quelli predisposti dagli apparati ministeriali, per monitorare l'evoluzione di una vasta gamma di fenomeni socio-economici (lavoro, devianza, evoluzione della popolazione, istruzione, salute, ecc.) e dare supporto al *policy-making*: dati che portano con sé non solo i difetti di un processamento non sempre professionale, ma anche la parzialità e i limiti strutturali della finalità istituzionale di origine (Zajczyk, 1994).

La tendenza alla datificazione delle notizie, in una prospettiva di oggettivazione del resoconto e dell'analisi della realtà sociale, viene sostenuta da un atteggiamento che diventa senso comune professionale (e sociale), e si estende ai sondaggi, quella forma di analisi demoscopica che dagli anni '30 del Novecento si afferma e diffonde nelle società democratiche e industriali avanzate, nella fase in cui l'opinione pubblica appare sempre meno comprensibile poiché attraversata dal processo di deideologizzazione (Corbetta e Gasperoni, 2007). Anche su tale fronte, i partecipanti allo spazio pubblico riversano fiducia nel dato statistico, che si ritiene riveli le tipiche fluttuazioni dell'incostante opinione pubblica di massa (Zaller, 1992), soprattutto nel rapporto con la politica. Tuttora si evidenzia una notevole resistenza a considerarne pienamente i limiti e le controindicazioni segnalate dal mondo accademico (Champagne, 1990; Natale, 2009; Di Franco, 2018). Dal lato del giornalismo, ciò richiama il combinato disposto dell'inerzia delle procedure e dei formati, e dello scarso livello di formazione sociostatistica degli operatori dell'informazione, specialmente nei contesti in cui l'accesso alla professione non sia – per legge o per consuetudine – connesso ad adeguati requisiti di competenza.

Per la professione giornalistica, l'insieme delle problematiche relative all'obiettività ha significato un incessante conflitto – prevalentemente interno ma anche esterno – tra visioni di ciò che debba essere la natura legittima della funzione e della prassi giornalistica, nel corso del quale il resoconto *fact oriented* e quello *data based* hanno rappresentato il contraltare di quello basato su atteggiamenti militanti (Hallin e Mancini, 2004; Kuypers, 2014) o di quello impostato sulla narrazione d'impronta letteraria (Hartscock, 2000; Sapiro, 2003). Tali configurazioni hanno costantemente costituito le polarità di dibattito e contrapposizione, investendo il giornalismo partigiano, quello investigativo, poi il cosiddetto *civic journalism*, e, all'opposto il giornalismo commerciale fino a sollevare in vario modo la cruciale questione dell'autonomia.

La difesa dell'indipendenza culturale e operativa del giornalismo è stata "aiutata" dall'avvento di Internet, ed esattamente dalle opportunità offerte dallo sviluppo delle pratiche di trasparenza proprie della *e-governance*, ossia di *openness* dei dati e delle informazioni, da parte sia di istituzioni ai

vari livelli, che di entità associative del vasto movimento internazionale della *open knowledge* (Ludovisi, 2016; Porlezza, 2016). Tale significativa trasformazione ha sollecitato un parallelo sviluppo di una figura ibrida di giornalista, che collabora con gli informatici o comunque possiede competenze di *data mining*, fino a irrobustire e legittimare ulteriormente il giornalismo di *precisione* (Meyer, 1973, 2004), sebbene questo sia stato assimilato con intensità diverse nei vari paesi occidentali (Ludovisi, 2016).

Più di recente è tuttavia emersa con maggiore forza la problematica della tecnologia e dei processi sociotecnici che presiedono alla produzione di dati e al loro prelievo-utilizzo da parte delle professioni che operano nello spazio pubblico, e più profondamente della datificazione della vita sociale generata dalla pervasività delle piattaforme (Lupton, 2014; van Dijck *et al.*, 2018). I processi di datificazione e di automazione algoritmica rappresentano oggi fattori rilevanti di riconfigurazione del campo giornalistico, capaci di incidere sulle pratiche professionali, sui criteri di produzione e validazione dell'informazione e sugli assetti decisionali delle organizzazioni mediatiche (Valentini, 2025).

Vi sono peraltro posizioni che vedono una linea di continuità tra tendenze alla trasparenza e l'attuale piattaformaizzazione delle modalità di estrazione ed elaborazione dei dati. Più di trent'anni fa, Power (1994) descriveva questo mutamento attraverso il concetto di *audit society*, riferendosi a una fase delle società industriali avanzate in cui la prominenza delle procedure di misurazione, tracciamento e verifica numerica acquisiscono un ruolo sempre più marcato nei processi di regolazione sociale.

Si tratta di una metamorfosi profonda che è venuta interessando le forme di *governance*, in un contesto di crescente complessità sociale. Il ricorso a principi quali trasparenza e *accountability* è stato infatti sempre di più realizzato attraverso una sistematica *valutazione* dell'azione politica ed economica fondata su risultanze quantitative (Mau, 2020). Con l'avanzare del XX secolo, tuttavia, è diventato maggiormente evidente come tali dinamiche eccedano l'ambito istituzionale, investendo più in profondità il tessuto della vita sociale, attraverso la diffusione capillare di dispositivi sociotecnici ibridi che partecipano attivamente ai processi di costruzione della realtà sociale (Lupton, 2014; Mennicken e Espeland, 2019; Mau, 2020; Bennato, 2024), arrivando a espandere ulteriormente lo spazio della datificazione.

In questo contesto, la mediazione giornalistica rischia di perdere la sua funzione, mentre invece diventa ancora più rilevante il suo ruolo nei processi di mediazione simbolica e discorsiva tramite i quali si negoziano e stabilizzano gli usi socialmente legittimi dei dati nello spazio pubblico.

Nel dibattito contemporaneo, segnato dalla circolazione di false notizie,

dall'opacità di alcuni sistemi di intelligenza artificiale e dall'uso propagandistico delle tecnologie digitali, è cresciuta l'attenzione verso il ruolo che il giornalismo può svolgere come «bussola sociale» (Sorrentino, 2025). Il *data journalism* viene spesso presentato come una risposta possibile: un rafforzamento di una vocazione antica del giornalismo, fondata sull'impiego dei dati – lo abbiamo detto in precedenza – non tanto e non solo come supporto alle notizie, quanto, più radicalmente, come risorsa per la costruzione e legittimazione dell'informazione.

La letteratura recente parla, a questo proposito, di una vera e propria «svolta quantitativa» (Petre, 2013), per indicare come i dati – nelle molteplici forme in cui possono essere raccolti, analizzati e presentati – siano diventati un imperativo per la professione giornalistica. Tale imperativo si è consolidato per almeno tre ordini di ragioni: (i) una spinta socio-politica, connessa agli ideali di trasparenza, *accountability* e controllo pubblico del potere; (ii) una spinta interna al campo giornalistico, legata all'ideale del giornalismo *watchdog* e alla sua affinità con pratiche di inchiesta basate su dati e analisi statistiche; (iii) una spinta infrastrutturale, derivante dalla pervasività di dispositivi socio-tecnici (tra cui i software di *data mining*) che generano quotidianamente volumi crescenti di dati. Ne consegue che il giornalismo opera oggi in una condizione paradossale: da un lato, è chiamato ad assumere il ruolo di selettore e interprete di fenomeni complessi per un pubblico generalista; dall'altro, si confronta con nuove forme di autorità epistemica, fondate su procedure statistico-computazionali, spesso opache, che competono con le modalità tradizionali di validazione giornalistica.

In tale configurazione del legame tra informazione e dati si gioca la pretesa di corrispondenza – istituzionalizzatasi nel secolo scorso – tra verità e certezza, ossia tra perfetta aderenza delle forme di rappresentazione alla realtà materiale dei fatti della vita sociale e capacità della scienza di fornire risorse di fuoriuscita dall'incertezza generata dal confronto con la complessità.

2. Genealogia del *data journalism*

Già oltre un secolo fa Walter Lippmann, in *Public Opinion* (1922) – testo comunemente considerato un punto di avvio nella riflessione sistematica sul rapporto tra giornalismo e opinione pubblica – attribuiva ai dati una funzione definitivamente risolutiva nella difficile missione sociale di ricerca della “verità”; soluzione efficace anche nel garantire sia l'obiettività sia l'autonomia della professione giornalistica in un mondo in cui irrompeva l'interdipendenza tra continenti (Marini, 2014).

Il suo ragionamento nasceva sullo sfondo del forte condizionamento e-

esercitato sui mezzi d'informazione dai vertici militari e dai governi nel contesto bellico e proseguiva nell'analisi delle pressioni che, in tempo di pace, avrebbero minacciato l'autonomia del giornalismo attraverso l'azione delle *public relations* e la distorsione strategica dell'informazione autoprodotta (nel segno della *publicity*) da parte di grandi aggregati economici e gruppi d'interesse.

A fronte di tali rischi, Lippmann proponeva (con una certa capacità predittiva) una soluzione coerente con un'impostazione di matrice positivista: la diffusione capillare di istituti e uffici di ricerca, dalle associazioni locali agli enti governativi, incaricati di produrre e rendere disponibili dati statistici su varie aree della realtà sociale. In questa cornice, la «verità dei fatti» appariva come un orizzonte raggiungibile attraverso l'ancoraggio della notizia a evidenze empiriche facilmente attingibili, che avrebbero innervato il discorso pubblico in modo sistemico, anche se Lippmann riconosceva l'im maturità delle scienze sociali rispetto al modello delle scienze «dure».

Si tratta quindi di una difesa dell'obiettività, contro l'influenza sovrachiantante degli stereotipi e contro l'eccesso di frammentazione dell'opinione pubblica; obiettività e verità che però il giornalismo realizza per merito della diffusione della conoscenza scientifica, delegando a questa una parte considerevole della propria funzione di mediazione culturale (Marini, 2014).

L'avvento dei sondaggi negli anni '30 e la fondazione delle società demoscopiche negli Stati Uniti realizzano un sistema di rilevazione ricorrente dell'opinione pubblica che da quel momento diviene tipico delle società democratiche (Corbetta e Gasperoni, 2007). Per il giornalismo tale sistema rappresenta un partner che rafforza la capacità delle testate d'interloquire efficacemente con la politica e con i centri di potere, avvalendosi di ciò che «manifesta» gli atteggiamenti della cittadinanza. I dati di sondaggio diventano stabilmente un indispensabile punto di partenza per la partecipazione di giornalisti, politici e consulenti (e degli stessi sondaggisti) a un processo di valutazione, previsione e discussione sulle cosiddette «opinioni», aspetti sfuggenti nel grande repertorio degli atteggiamenti di cui sarebbe portatrice la fluttuante opinione pubblica, intesa non più come soggetto di *agency*, ma come aggregato indistinto e incostante (Price, 1992; Zaller, 1992). Si rende stabile il monitoraggio di variabili quali l'approvazione e la fiducia verso i leader politici (specialmente presidenti e premier), il problema più urgente da risolvere, la fiducia negli sviluppi futuri da parte di imprese e consumatori, le intenzioni di voto e anche opinioni su aspetti contingenti ma controversi delle decisioni politiche in formazione.

Nei decenni successivi, tale struttura si impianta anche nei paesi europei, dando spazio a una sorta di generalizzata ingegnerizzazione del rappor-

to tra cittadini e politica, con il trasferimento di tali tecniche e lo sviluppo di professioni specializzate (Nimmo, 1970; Sabato, 1981; Champagne, 1988, 1990). Ciò alimenta una comunità interpretativa di professionisti vari, che istituzionalizza a fini sia strategici che comunicativi l'uso dei dati di sondaggio, ritenuti capaci di rilevare (e rivelare) oggettivamente lo stato dell'opinione pubblica, grazie a standard codificati di elaborazione. Questo sviluppo apre le porte anche alla consuetudine di rilanciare dati, di varia provenienza e non necessariamente organizzati, nel discorso pubblico per la gestione comunicativa delle contingenze.

Dopo l'avvento dei sondaggi, a partire dagli anni '40, l'ideale dell'obiettività viene però investito da una controversia che attraversa il mondo giornalistico e politico-culturale. Cresce infatti la critica all'obiettività come «feticcio», e, con essa, all'idea che il giornalismo possa offrire un resoconto efficace del reale e soprattutto utile alla crescita democratica (Siebert *et al.*, 1956; Nerone, 1995; Marini, 2021). In questa fase si afferma una diversa attribuzione di funzione al giornalismo, che trova una formulazione influente in una delle quattro teorie di Siebert e colleghi, quella della *responsabilità sociale*: non più la «verità dei fatti», ma la «verità intorno ai fatti». La trasformazione si regge almeno su due pilastri: da un lato, la legittimazione di un giornalismo interpretativo (a lungo osteggiato nel contesto nordamericano: Blanchard, 1977); dall'altro, la concezione dell'informazione come infrastruttura di un forum pubblico di dibattito, aperto alle forze sociali, anche a quelle minoritarie.

Lo spostamento del focus dal fatto al suo contesto rappresenta l'assunzione di una prospettiva di riconnessione delle notizie alle sequenze e regolarità che rendono i fatti classificabili e intellegibili. Come diremo più avanti, è una prospettiva che rende possibile e necessario il collegamento tra eventi, fenomeni e temi, che non esclude il ricorso a dati statistici, come coadiuvante dell'analisi e dell'interpretazione. Vi si legge anche l'idea di un'alleanza con la società civile, grazie alla quale il giornalismo acquista la capacità, ma anche viene sollecitato ad essere sensibile alle condizioni della vita sociale e ai *claims* da essa emergenti (Marini, 2021).

Tale spostamento si collega rapidamente alla cultura dei *social problems*, radicata nella postura progressista, in particolare alle riflessioni di Herbert Blumer (1971) sui problemi sociali come prodotto di processi collettivi di riconoscimento e definizione. Blumer contesta l'esistenza "oggettiva" (mostrata dai dati) di condizioni sociali che manifestano problemi, avendo quindi una natura intrinseca, rilevabile con il confronto tra funzionale e disfunzionale, tra normale e deviante (ivi). Piuttosto, per Blumer i problemi sono esiti di processi di costruzione sociale: una visione antipositivistica che sottolinea la capacità congiunta di esperti, attivisti e giornalisti

di rendere visibili criticità, farle emergere come questioni socialmente rilevanti e reclamare azioni politiche, come affermato nella teoria dell'*agenda building* e nel consolidarsi dell'approccio costruzionista ai problemi sociali (Marini, 2006; Best, 2013).

Questo tema si colloca alla base di una serie di studi di tipo quantitativo che si pongono l'obiettivo di analizzare la coerenza o la discrepanza tra l'evoluzione dei fenomeni sociali e la sensibilità dell'opinione pubblica in termini di preoccupazione-urgenza, confrontando i dati "oggettivi" di analisi statistica istituzionale con i risultati di sondaggio in serie storica (Funkhouser, 1973; Baumgartner e Jones, 1993). La conclusione generale è che il *public concern* mostra andamenti disallineati dai cicli della cosiddetta «vita reale», poiché seguono invece la ciclicità dell'attenzione pubblica, che viene mobilitata, anche a seguito di eventi straordinari, da parte di media, movimenti sociali e gruppi d'interesse di vario tipo (Dearing e Rogers, 1996; Jones e Baumgartner, 2005).

Riprendendo il filo storico degli anni '70, una risposta contrastiva a Blumer e alla sua impostazione costruzionista viene di lì a poco da Morris Janowitz (1975), che propone una classificazione (non casualmente dicotomica) dei modelli professionali di giornalismo, contrapponendo il *gatekeeper* all'*advocate*. Il primo è orientato alla ricerca dell'obiettività e segue la regola della «netta separazione» tra resoconto dei fatti e disseminazione delle opinioni; il secondo patrocina le istanze di gruppi sociali a lui affini culturalmente e politicamente, assimilandosi alla professione legale. Qui il *gatekeeping* è concepito solo come selezione di eventi che possano interessare il pubblico di massa (o soddisfare le finalità economiche dei gruppi editoriali): viene rifiutata l'idea che il giornalismo selezioni eventi notiziabili avendo in precedenza identificato – attraverso le relazioni e le interdipendenze con altri attori e “forze” sociali – temi rilevanti per poterli articolare e discutere, anche grazie all'utilizzo di dati. Questa risorsa pertanto – nella versione della responsabilità sociale e dell'*advocacy* –, seppure pubblicamente disponibile, non si colloca a monte ma a valle o in parallelo con l'emersione delle *issues*; i dati possono essere prelevati e a loro volta selezionati, nonché rielaborati, a seconda della prospettiva da cui il tema-problema venga analizzato. Analogamente, si può dire che nella linea ideologica che assegna il primato ai fatti e ai dati, al giornalismo non si attribuisce, e forse si nega, la prerogativa di possedere e promuovere visioni del mondo e di realizzare proprie strategie di tematizzazione.

In questo ampio arco temporale si determina quindi una frattura interna al mondo giornalistico e – lo si può solo accennare – un distanziamento tra diverse esperienze nazionali e continentali, in linea di massima tra “modello” americano e “modello” europeo (Hallin e Mancini, 2004). Tale *cleavage*

vede, da un lato, oggettivismo e idealizzazione della neutralità informativa, e, dall'altro, interpretazione e funzione "deliberativa" dell'informazione.

Tale tensione – sullo sfondo dell'antagonismo tra correnti politico-culturali – rende anche conto delle ulteriori trasformazioni negli orientamenti e nelle pratiche della professione che hanno portato ad affermare il giornalismo dei dati (*data journalism*). Non a caso, gli anni '60 negli Stati Uniti sono il contesto spazio-temporale di un cambiamento del giornalismo, che cerca di recepire spinte sociali ed economiche diverse, nel segno dell'innovazione. Mentre si afferma un filone orientato verso il racconto *fictional* degli eventi (*new journalism*), un indirizzo opposto tenta la strada di un'autonomia professionale che sancisca l'abbandono di ogni acquiescenza verso il potere, anche come esito di una rivendicazione generazionale da parte di giovani giornalisti (Schudson, 1978). Tra gli strumenti professionali viene privilegiata l'inchiesta, corrispondente a un'inclinazione, fino ad allora minoritaria, verso l'approfondimento critico. All'interno di tale linea, emerge il giornalismo «di precisione», che spinge in alto gli standard del rigore e dell'accuratezza, attraverso l'analisi puntuale dei documenti, e attraverso ricerche realizzate con l'adozione delle procedure degli scienziati sociali. I reportage privilegiano l'uso di dati al posto delle tradizionali interviste e testimonianze, effettuando in autonomia sondaggi, o lavorando su *dataset* istituzionali. Si tratta di giornalisti che intendono applicare il metodo delle scienze statistico-sociali (e in seguito con l'assistenza delle competenze informatiche), al fine di mettere in luce contraddizioni e ingiustizie, derivanti ad esempio dalle discriminazioni razziali (Meyer, 1973; Gandy, 1998).

3. Il *data journalism* tra conoscenza aperta e controversia interpretativa

Grazie a questa versione del *data journalism*, la frattura di cui abbiamo parlato prima tende a mitigarsi, poiché prassi e culture professionali in trasformazione avvicinano le opposte polarità. Il ricorso ai dati viene sempre più spesso incorporato in pratiche che li prelevano selettivamente in modo finalizzato, li contestualizzano e li sottopongono a interpretazione, pur nei limiti e nelle variazioni prodotte dall'evoluzione del contesto nelle diverse aree politico-culturali (Esser e Umbricht, 2013; Ludovisi, 2016).

In quella che appare come una variegata pluralità di esperienze e orientamenti, rimane in discussione, anche se non apertamente, il ruolo *dirimente* dei dati. Ha infatti peso diverso l'eredità dell'oggettivismo, ossia la misura variabile in cui i differenti contesti e culture professionali l'hanno assorbita (o recuperata tardivamente). Tale assimilazione, tuttavia, non può esse-

re letta come trasformazione realizzata solo dall'interno della professione, poiché riguarda – come detto – cambiamenti negli ecosistemi informativi, in cui interagiscono diversi attori e logiche, travalicando il perimetro del mondo giornalistico.

Il giornalismo dei dati (o con i dati), infatti, non avrebbe potuto trovare spazio operativo se non si fosse determinato lo sviluppo (in qualche modo simile a quello auspicato da Lippmann) della produzione e pubblicazione di *database* da parte – come detto all'inizio – di istituzioni, enti pubblici e privati, rispondenti a una crescente domanda di trasparenza da parte degli *stakeholders* (anche in tal caso nel segno della responsabilità sociale). I cosiddetti *open data* sono il prodotto della tendenza alla *open knowledge*, che coinvolge anche le istituzioni rappresentative, come nei numerosi casi di siti governativi e parlamentari creati – anche in Italia – per la “pubblicità” di dati sulle proprie attività (Ludovisi, 2016). Questo trend ha sollecitato la formazione di organizzazioni *non-profit* e indipendenti, nonché di loro network internazionali: finalizzate alla promozione della trasparenza e alla raccolta-distribuzione di dati di rilevanza pubblica (ivi).

La sua finalità è connessa a un ideale di partecipazione, risalente al cyber-ottimismo del primo decennio di Internet. Si prefigge cioè non solo di offrire strumenti di partecipazione consapevole alla società civile e alla cittadinanza in generale, ma anche di ottenere feedback interattivi in funzione di controllo dello stesso operato dei produttori-distributori di dati e informazioni (ivi). In questa prospettiva, il giornalismo dei dati opera interagendo con tutti gli attori *stakeholders* della trasparenza, secondo un modello allargato di *e-governance*, che però non può che essere riferito a pubblici specializzati, sebbene pratiche di interazione cooperativa mostrino notevoli potenzialità (Batsell, 2015).

Occorre aggiungere che la diffusione delle attività di partecipazione al dibattito pubblico si è accompagnata alla consapevolezza del carattere strategico della produzione di dati statistici. Ciò ha portato molti attori dello spazio pubblico, specialmente quelli appartenenti alla società civile, quali organizzazioni di rappresentanza degli interessi e movimenti o grandi associazioni, a sviluppare la capacità di produrre, pubblicare e promuovere ricerche, dossier statistici, report ricorrenti e osservatori. Esiste una pluralità di enti non istituzionali che genera una vasta produzione di dati, che si affianca e spesso compete con quella istituzionale. Si possono citare i report di Greenpeace sui rifiuti industriali, di Amnesty International sulla pena di morte e sui prigionieri politici, di Freedom House sulla libertà d'informazione; a livello nazionale quelli di Legambiente sull'inquinamento delle acque marine e interne, e della Caritas sull'immigrazione. Tutti sono il risultato di collaborazioni tra attivisti, istituzioni accademiche ed esperti di vario

tipo, ed anche di una crescente professionalizzazione della comunicazione a fini di *news management*.

Si tratta del tentativo non solo di instaurare relazioni stabili con le testate d'informazione, ma anche di affermare uno *standing* nel settore ed esercitare influenza. Sulla stessa linea si muovono vari enti privati o semipubblici, soprattutto fondazioni, al cui interno operano uffici-studio, spesso definiti *think tank*, che hanno lo scopo di occupare uno spazio di produzione e diffusione di conoscenza *data based* su nicchie tematiche. I media italiani hanno collaborato frequentemente negli ultimi anni con enti di questo tipo: ad esempio con la Fondazione Gimbe in epoca di Covid-19, per il monitoraggio della pandemia e dell'efficacia delle campagne vaccinali; con Nomisma Energia per l'andamento del mercato globale degli idrocarburi e connesse previsioni, nelle due crisi energetiche recenti.

Gli ambienti comunicativi sono notevolmente datificati, e lo spazio pubblico si è quindi addensato di tali forme di produzione strategica di statistiche e di report: questi sono divenuti una risorsa discorsiva per la partecipazione al dibattito pubblico, a partire dalla cattura dell'attenzione dei media, dei pubblici sensibili e della politica. Pertanto, la mediazione giornalistica si muove all'interno di un sistema di interdipendenze che ne pone spesso in discussione la terzietà (Sorrentino, 2008), poiché il giornalismo deve mettere in campo competenze e capacità di adattamento molto avanzate (Antenore e Splendore, 2017), difficili da dispiegare, soprattutto nel corso della crisi che la transizione digitale comporta (Ludovisi, 2016).

La prassi giornalistica entra infatti in interazione con un più ampio ecosistema di produzione della conoscenza che riverbera su di essa le dinamiche competitive o anche conflittuali per la definizione della rilevanza. L'uso dei dati può essere ricondotto a tre dimensioni – e, al contempo, a tre momenti del lavoro giornalistico – che strutturano il passaggio dalla descrizione alla costruzione pubblica del significato: eventi, fenomeni sociali e temi. I fenomeni sociali possono essere intesi come aggregati di eventi, ma la loro stessa aggregazione presuppone un lavoro interpretativo: gli eventi vengono selezionati, ordinati e stabilizzati entro categorie ed etichette linguistiche condivise, tanto nelle comunità scientifiche quanto nelle comunità interpretative di cui i giornalisti fanno parte.

Il passaggio più rilevante è tuttavia quello dai fenomeni ai temi, cioè a quegli aggregati discorsivi che consentono di ragionare su eventi e fenomeni attraverso percorsi di *framing* e chiavi di lettura, generalmente ancorati a orientamenti valoriali e visioni del mondo, che dai temi vengono attualizzate.

In questa prospettiva, i giornalisti non “applicano” semplicemente dati a una realtà già configurata, ma partecipano a processi collettivi di interpretazione e problematizzazione. Negli ambienti professionali in cui si è svilup-

pata competenza e distacco dall'oggettivismo dei dati, i giornalisti possono utilizzare i dati in modo finalizzato, possono prendere distanza dalla presunta oggettività della misurazione e rendere visibile il carattere situato delle rappresentazioni. Proprio per questo, come accade per i temi, anche l'uso dei dati tende a generare controversie: la quantificazione non neutralizza il conflitto interpretativo, ma può divenirne uno degli oggetti privilegiati.

Ciò si collega al fatto che i fenomeni e i temi hanno intrinsecamente una sostanza multidimensionale, che i set di dati utilizzati inevitabilmente riducono, rendendo parziale – e fuorviante – l'operazione di chiarificazione attesa dall'uso stesso dei dati.

La prassi giornalistica tende a recuperare la multidimensionalità con un atteggiamento *multiprospettico*, realizzato talvolta con l'analisi di altri dati, ma più spesso per mezzo di altri tipi di narrazione. Una possibilità è la ricognizione e raccolta delle prese di posizione e delle interpretazioni da parte di una pluralità di attori (Benson, 2009; Marini e Gerli, 2017): pratica che risponde ai classici criteri di pluralismo e imparzialità, ma talvolta esprime la natura rituale dell'adozione di una regola, anch'essa tendente ad esaudire le richieste di obiettività (Tuchman, 1972). Un'altra possibilità è di spostarsi sulla narrazione di *human interest*, raccontando storie personali di protagonisti o testimoni, prelevate sul campo (Marini e Gerli, 2023). Del resto, questo è notoriamente il carattere tipico delle tematizzazioni, che vengono articolate tramite *shifting* semantici e di genere; cosa che riconcilia quei tipi di giornalismo che la rigidità dell'obiettivismo ha a lungo considerato incompatibili.

Nel solco delle riflessioni sviluppate da Pierre Bourdieu (1992, 1996a, 1996b; Gerli, 2024), il rapporto tra rappresentazione giornalistica e realtà sociale può essere inteso come una *posta in gioco*, vale a dire come un processo intrinsecamente controverso, negoziale e dinamico, suscettibile di stabilizzazioni temporanee ma costantemente alimentato dal conflitto interpretativo. Il giornalismo si configura così come un campo plurale e conflittuale, dotato di un'autonomia sempre relativa rispetto alle proprie regole del gioco e alle proprie procedure (Gerli, 2024), ma al tempo stesso strutturalmente esposto alla negoziazione continua con una molteplicità di istanze esterne – dalla politica all'economia, dalla società civile organizzata alla cultura e alla sfera religiosa. È precisamente questa *configurazione relazionale* che rende inevitabile la diversificazione delle pratiche *data-driven* e che colloca l'uso dei dati all'interno di una più ampia dinamica di conflitti per la definizione delle rappresentazioni legittime della realtà sociale.

Sul piano pratico, questa pluralità si traduce in modalità differenziate di utilizzo dei dati nel giornalismo, riconducibili all'autonomia relativa del campo, alle pressioni sistemiche e alle pratiche situate delle singole redazioni, in relazione alla loro strategia editoriale e alla specificità degli eventi

trattati. Ne consegue che il dato non opera come evidenza neutrale, ma come risorsa simbolica mobilitata all'interno di strategie concorrenti di rappresentazione e posizionamento (Fourcade e Healy, 2017; Mau, 2020), la cui osservazione empirica consente di mettere a fuoco i meccanismi attraverso cui il giornalismo contribuisce a definire pubblicamente il senso dei fenomeni sociali.

Il sondaggio d'opinione – come vedremo nel seguente paragrafo – costituisce in questo senso un dispositivo emblematico di tale dinamica: dietro l'apparente neutralità statistica, esso produce un effetto di democraticità che occulta un intervento diretto nella costruzione e nella gerarchizzazione dei temi politicamente rilevanti (Bourdieu, 1973; Champagne, 1988, 1990; Gerli, 2026); inoltre, esso offre un sufficiente grado di regolarità operativa nei processi di produzione di notizie, sempre orientati alla routinizzazione (Epstein, 1973; Tuchman, 1978). Una volta integrati nel circuito mediatico, i sondaggi d'opinione cessano di essere meri strumenti di rilevazione e diventano risorse simboliche mobilitabili nella competizione politico-culturale, nella misura in cui contribuiscono non solo a misurare, ma anche a selezionare, gerarchizzare e legittimare determinati temi e rappresentazioni della realtà sociale (Champagne, 1990). In questo senso, la loro circolazione conferisce al giornalismo un potere specifico nella definizione dello spazio del dicibile e del pensabile, trasformando la quantificazione dell'opinione in un elemento centrale della competizione per l'interpretazione dei fenomeni sociali.

4. La costruzione *data-driven* del tema immigrazione nel giornalismo italiano

La copertura giornalistica dell'immigrazione in Italia offre un terreno empirico particolarmente fertile per osservare il funzionamento di queste dinamiche. Tema strutturalmente controverso e fortemente esposto alla competizione tra attori politici, istituzionali, mediali e della società civile organizzata, l'immigrazione costituisce un caso in cui i dati vengono frequentemente mobilitati per costruire, legittimare o contestare specifiche rappresentazioni della realtà sociale.

In questa sezione viene quindi problematizzata la tendenza ad assumere i numeri – e i loro derivati, quali grafici e visualizzazioni – come risorse discorsive che producono *effetti* di oggettività, contribuendo a ridurre la visibilità delle negoziazioni interpretative che ne accompagnano la costruzione e l'uso. A supporto dell'argomentazione, vengono presi in esame alcuni esempi tratti dal giornalismo italiano contemporaneo, relativi alla rappresentazione *data-driven* dei fenomeni migratori.

Il dibattito pubblico sull'immigrazione in Italia è sempre stato caratterizzato da controversie ricorrenti sui dati. In generale, tali diatribe hanno spesso mostrato un uso parziale e tendenzioso delle statistiche e dei risultati di ricerca, al quale il giornalismo non si è sottratto, venendo trascinato dalla conflittualità politico-culturale (Sciortino e Colombo, 2004; Diamanti, 2017; Marini e Gerli, 2017) su una delle questioni che maggiormente hanno ridisegnato il panorama delle culture politiche a livello internazionale e in particolare europeo.

Tramite la loro apparente neutralità, anche ai giornalisti i dati appaiono dirimenti nel contesto di una circolazione di tesi politiche e di notizie destinata a creare confusione. Si evidenzia un contrasto tra due tipi d'informazione: quella di cronaca, soprattutto locale, che si nutre degli eventi di devianza e criminalità, delle crisi dei centri di accoglienza, delle proteste di quartiere ecc.; quella di approfondimento, incline a evitare il sensazionalismo ma inevitabilmente meno frequentata. Ciò sembra dare luogo a una contraddizione anche in termini di linguaggio giornalistico difficilmente sanabile (Marini *et al.*, 2019). La frequente commistione tra cronaca e dati sembra mostrare un'assimilazione del dato come dispositivo narrativo di supporto, come richiamo rituale alla certezza numerica.

Tale dinamica si è svolta in parallelo con l'incremento dell'impegno da parte di entità del Terzo settore (Cestim, Ismu, Caritas, Idos) nella produzione di statistiche e indagini. Questo conferma il modello di partecipazione di cui abbiamo parlato in precedenza, e in questo caso vi aggiunge l'aspetto strategico di aperto contrasto alla xenofobia da parte di un'*advocacy coalition* che ha spesso coinvolto i mezzi d'informazione liberal-progressisti (Marini, 2021). Questo tipo di coalizione, tuttavia, si è formata e attivata anche sul fronte opposto, con un network di associazioni, think tank (come Nazione Futura e Centro Machiavelli), case editrici, periodici e quotidiani di orientamento nazionalista e anti-immigrazione, che si occupa prevalentemente di promozione culturale e politica ma anche di divulgazione inchieste e reportage, finalizzati a contrastare le varie espressioni della *correttezza politica* delle sinistre.

I nuclei tematici di utilizzo dei dati su cui la controversia si è più spesso animata nel contesto mediale italiano sono stati le presenze sul territorio nazionale e la devianza-criminalità: si tratta di sottodimensioni della xenofobia piuttosto elementari, e in quanto tali molto popolari e adatte ad agitare l'informazione generalista. Sono questioni strettamente collegate al *cleavage* politico-culturale accettazione *vs* rifiuto, i rami principali di una tassonomia tematica che va ad articolarsi in modo esteso, seguendo i due grandi frame della solidarietà e della minaccia.

Riguardo alla quantificazione delle presenze, una pietra miliare del mo-

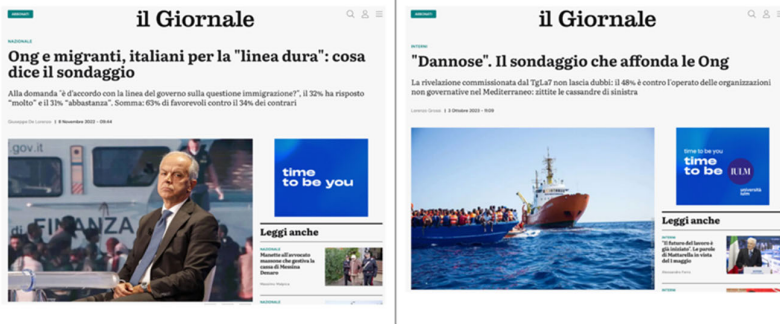
dello d'intervento di tipo *advocacy* è il *Dossier statistico immigrazione*, pubblicato da Caritas-Migrantes per la prima volta nel 1991 (dopo la prima importante legge per la regolarizzazione) proprio per fare chiarezza sulle presenze, ripulendo i dati spuri elaborati dal Ministero dell'Interno; per proseguire, fino a oggi, la sua missione di contrasto ai pregiudizi e alla scorretta informazione tramite la statistica. Una controinformazione statistica che, ancora una volta, ha come controparte – diciamo così – l'effetto matematico indiretto della cronaca giornalistica.

La seconda questione, quella della criminalità, si è sviluppata in modo ancora più densamente caratterizzato da controversie mediatiche, alimentate dal *bias* sopra menzionato tra due tipi di produzione giornalistica, a fronte di una limitata e spesso poco gradita produzione scientifica (Barbagli, 1998, 2008). Il trend di incremento dell'incidenza della devianza straniera a partire dagli anni '90 ha generato molteplici polemiche sui dati, specialmente intorno alle interpretazioni pretestuose e banalizzanti sull'associazione *tout court* tra criminalità e immigrazione (testimoniate anche nella ricerca sui media: si veda Binotto *et al.*, 2016). Si sono poi manifestati segnali di un assorbimento normalizzante del tema nella produzione giornalistica, meno sensibile alla contrapposizione politica polarizzata, anche perché superato da altre “emergenze” e controbilanciato da altre questioni (Marini *et al.*, 2019); il che sembra corrispondere alla valutazione più distaccata dei dati sul fronte della società civile (ad es. Caritas Migrantes, 2017). Tale tendenza ha visto una parallela maggiore visibilità pubblica di indagini d'ispirazione populista sulla criminalità dei migranti (Scalea, 2019; Totolo, 2024) riecheggiate dal coordinamento tra media e centri-studi di destra.

Gli esempi relativi alla copertura giornalistica dell'immigrazione, che proponiamo qui di seguito, offrono un terreno empirico utile per osservare il funzionamento del meccanismo di uso dei dati in funzione della controversia e, oppure, in funzione di un arbitrato statistico per il suo superamento. In diversi articoli basati su sondaggi realizzati da istituti privati – spesso attivi anche nel settore delle ricerche di mercato –, le percentuali vengono presentate come misure oggettive dell'opinione pubblica nel contesto del conflitto politico legato al cosiddetto “decreto Piantadosi” e, più in generale, alla governance dei flussi migratori e al ruolo delle ONG nel Mediterraneo. Titoli che enfatizzano una presunta maggioranza favorevole alla “linea dura” o alla chiusura delle frontiere traducono risposte individuali aggregate in rappresentazioni compatte del sentire collettivo, inscrivendole in frame polarizzati (Fig. 1). Anche quando prodotti da istituti riconosciuti, tali dati vengono mobilitati attraverso domande che orientano l'interpretazione (ad esempio associando l'immigrazione al pericolo per l'ordine pubblico) e successivamente ricondotti a schemi oppositivi fondati sull'orientamento

politico dei rispondenti, rafforzando dicotomie tra cultura dell'accoglienza e cultura delle restrizioni (Fig. 2).

Fig. 1 - Sondaggi e immigrazione ne Il Giornale (nov. 2022 e ott. 2023)



Fonte: Il Giornale.

Fig. 2 - Sondaggi e immigrazione ne la Repubblica (sett. 2023 e mar. 2024)



Fonte: la Repubblica.

Accanto a queste modalità di impiego dei sondaggi, emergono anche pratiche giornalistiche che propongono letture "alternative" dei fenomeni migratori attraverso differenti selezioni di dati, fonti e cornici interpretative.

Non si tratta semplicemente di un rovesciamento ideologico delle posizioni politiche, ma dell'esito di un diverso assetto dei materiali empirici mobilitati e delle modalità con cui vengono organizzati e interpretati. Alcune inchieste *data-driven* offrono, in tal senso, una rappresentazione significativamente diversa del ruolo delle ONG rispetto a quella veicolata dai sondaggi d'opinione presentati come misurazioni oggettive del sentire collettivo. Un esempio significativo è rappresentato dal *Dataroom* di Milena Gabanelli (Fig. 4): basandosi prevalentemente su dati provenienti da fonti istituzionali (Ministero dell'Interno, Eurostat, Istat, OMS), l'inchiesta ricostruisce l'evoluzione delle politiche migratorie e delle dinamiche dei flussi nel Mediterraneo, mettendo in discussione l'associazione causale tra presenza delle ONG e aumento delle partenze dal Nord Africa ("Migranti: 10 anni di errori, ipocrisie, propaganda e il falso problema delle ONG"). In questo caso, i dati quantitativi vengono mobilitati per sostenere una lettura dei flussi migratori che relativizza il ruolo delle ONG nelle partenze dal Nord Africa, inserendo l'andamento degli sbarchi all'interno di una ricostruzione più ampia delle politiche migratorie europee e italiane, degli accordi internazionali e dei mutamenti istituzionali intervenuti nel tempo. Ne emerge una forma di giornalismo *data-driven* orientata alla storicizzazione e contestualizzazione dei fenomeni attraverso l'integrazione di serie storiche, dati istituzionali e comparazioni temporali, piuttosto che alla riduzione del fenomeno a spiegazioni semplificate o monocausali.

Fig. 3 - Sondaggi e immigrazione ne Il Giornale e Il Fatto Quotidiano (nov. 2024)



Fonte: il Giornale e il Fatto Quotidiano.

Analoghi dispositivi emergono in sondaggi utilizzati per legittimare posizioni critiche verso la magistratura o per consolidare letture emergenziali dei fenomeni migratori, mostrando come la quantificazione possa diventare

parte integrante di strategie discorsive di tipo politico (Fig. 3). In tutti questi casi, il sondaggio non si limita a registrare orientamenti preesistenti, ma contribuisce a *costruire* un’“opinione pubblica” apparentemente oggettiva, capace di conferire autorevolezza a specifiche interpretazioni dei fenomeni sociali e di orientare il dibattito mediatico e politico.

Fig. 4 - Politiche migratorie e andamento degli sbarchi nel Dataroom del Corriere della Sera (nov. 2022)

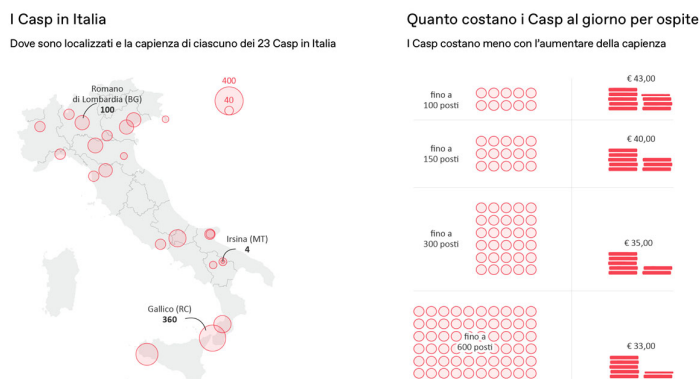


Fonte: Dataroom – Corriere della Sera.

Una funzione analoga può essere attribuita ad altri prodotti informativi che utilizzano il dato come strumento di riorientamento delle cornici inter-

pretative attraverso cui i fenomeni migratori vengono rappresentati nello spazio pubblico. Articoli come “L’accordo Italia-Albania, uno spot da 650 milioni” (*Corriere della Sera*, marzo 2024), costruiti prevalentemente su dati di origine istituzionale, propongono una lettura delle politiche migratorie centrata sui costi economici, sulle implicazioni amministrative e sull’efficacia delle misure adottate dal Governo italiano. In questi casi, la quantificazione viene mobilitata all’interno di una cornice interpretativa orientata all’analisi delle *policy* pubbliche e dei loro effetti.

Fig. 5 - CASP e sistema di accoglienza nella copertura data-driven di IRPI Media (lug. 2024)



Dati: Viminale 2024 • Sono inclusi costi per servizi di accoglienza, vitto, pulizia e prestazioni aggiuntive • Irpimedia 2024

Didier* è arrivato in Italia con una gamba rotta e tanti pensieri in testa. Colpa del difficile viaggio dalla Costa d’Avorio, delle violenze subite in Libia e della sorte della figlia con la quale ha perso i contatti, rimasta nel suo Paese d’origine. Avrebbe avuto bisogno di un luogo adatto alle sue condizioni fragili, sia a livello fisico sia a livello psicologico. E invece si è ritrovato in un **Casp**, il nuovo, ultimo e più basso gradino del sistema d’accoglienza italiano.

I Casp sono Centri di accoglienza straordinari provvisori e quello in cui è stato Didier per oltre un mese si trova a Martorano, in provincia di Parma. È un gruppo di container dentro un’ex fabbrica dismessa a 15 chilometri dalla città, in piena campagna. Lo gestisce la cooperativa sociale World in Progress che garantisce agli ospiti solo i servizi minimi.

Fonte: IRPI Media.

Una logica simile si ritrova anche in contributi pubblicati su *Info Data*, il data blog de *Il Sole 24 Ore*, nei quali grafici, tabelle e infografiche vengono utilizzati come dispositivi di organizzazione e selezione dell’informa-

zione, contribuendo a rendere intelligibili fenomeni complessi attraverso pratiche di quantificazione e comparazione. Titoli come “L’Italia non ha solo il problema della fuga di cervelli. Perché non è in grado di attirare immigrati laureati?” (luglio 2021), “Vale 154 miliardi, ovvero il 9% del totale, il prodotto interno lordo generato dai lavoratori immigrati in Italia” (ottobre 2023) e “Nuovi nati, decessi e migranti. Cosa è successo ai comuni italiani in 5 anni?” (gennaio 2023) mostrano come la quantificazione venga impiegata per tematizzare l’immigrazione in relazione a dinamiche strutturali del mercato del lavoro, dell’andamento demografico e della sostenibilità economica. Ulteriori esempi, provenienti dal giornalismo investigativo e indipendente (per esempio, *IRPI Media*), mostrano come il dato possa essere proficuamente combinato con testimonianze dirette e narrazioni *human-interest*, producendo rappresentazioni al tempo stesso empiricamente fondate e socialmente situate (Fig. 5).

Nel loro insieme, questi casi suggeriscono che il problema non risiede nei dati in quanto tali, ma nelle modalità attraverso cui vengono selezionati, interpretati e inseriti in dispositivi narrativi. Più che ricorrere alla categoria di *bias* – che presuppone un ideale di neutralità difficilmente sostenibile in un campo strutturalmente plurale e conflittuale – appare più produttivo distinguere tra pratiche giornalistiche come espressione dei diversi posizionamenti delle testate nello spazio sociale e delle configurazioni relazionali che si vengono a creare, di volta in volta, attorno a eventi specifici, coinvolgendo istituzioni, fonti, dispositivi di produzione dei dati e attori dello spazio pubblico. Esistono usi del dato orientati alla semplificazione e alla polarizzazione, così come pratiche capaci di attivare forme di riflessività pubblica e di interrogare criticamente i dispositivi di quantificazione. Tali differenze si intrecciano con le condizioni sociali di ricezione: alcuni prodotti informativi sono concepiti per pubblici con maggiore disponibilità all’approfondimento e alla complessità.

Riferimenti bibliografici

- Antenore M., Splendore S. (a cura di) (2017), *Data journalism. Guida essenziale alle notizie fatte con i numeri*, Mondadori, Milano.
- Barbagli M. (1998), *Immigrazione e criminalità in Italia*, Il Mulino, Bologna.
- Barbagli M. (2008), *Immigrazione e sicurezza in Italia*, Il Mulino, Bologna.
- Batsell J. (2015), *Engaged Journalism: Connecting with Digitally Empowered News Audiences*, Columbia University Press, New York.
- Baumgartner F.R., Jones B.D. (1993), *Agendas and Instability in American Politics*, Chicago University Press, Chicago.

- Bennato D. (2024), *La società del XXI secolo. Persone, dati, tecnologie*, Laterza, Roma-Bari.
- Benson R. (2009), "What makes news more multiperspectival? A field analysis", *Poetics*, 37(5-6), 402-418.
- Beretta M. (2017), *Storia materiale della scienza*, Carocci, Roma.
- Best J. (2013), "Constructionist Social Problems Theory", *Communication Yearbook*, 36, 237-269.
- Binotto M., Bruno M., Lai V. (a cura di) (2016), *Tracciare confini. L'immigrazione nei media italiani*, FrancoAngeli, Milano.
- Blanchard M.A. (1977), "The Hutchins Commission, the Press and the Responsibility Concept", *Journalism Monographs*, 49, 1-51.
- Blumer H. (1971), "Social Problems as Collective Behavior", *Social Problems*, 18, 298-306.
- Bourdieu P. (1973), « L'opinion publique n'existe pas », *Les Temps Modernes*, 318, 1292-1309.
- Bourdieu P. (1992), *Les règles de l'art. Genèse et structure du champ littéraire*, Seuil, Paris.
- Bourdieu P. (1996a), *Champ politique, champ des sciences sociales, champ journalistique*. 5° Cahiers de recherche Gps (Groupe de recherche sur la socialisation) de l'Université Lumière, Lyon.
- Bourdieu P. (1996b), *Sur la télévision, suivi de L'emprise du journalisme*, Liber, Paris.
- Caritas Migrantes (2017), *Dossier statistico immigrazione*, Idos, Roma.
- Champagne P. (1988), "Le cercle politique. Usages sociaux des sondages et nouvel espace politique", *Actes de la Recherche en Sciences Sociales*, 71-72, 71-97.
- Champagne P. (1990), *Faire l'opinion. Le nouveau jeu politique*, Éditions de Minuit, Paris; trad. it. M. Gerli (a cura di) (2026), *Fare l'opinione. Media, sondaggi e potere simbolico*, Armando, Roma (in corso di pubblicazione).
- Corbetta P., Gasperoni G. (2007), *I sondaggi politici nelle democrazie contemporanee*, Il Mulino, Bologna.
- Dearing J.W., Rogers E.M. (1996), *Agenda-Setting*, Sage, Thousand Oaks.
- Diamanti I. (2017), "Immigrati, cresce la paura: il 46% si sente in pericolo. È il dato più alto da dieci anni", *la Repubblica*, 13 settembre 2017; anche in Demos: <http://www.demos.it/a01428.php>.
- Di Franco G. (2018), *Usi e abusi dei sondaggi politico-elettorali in Italia. Una guida per giornalisti e ricercatori*, FrancoAngeli, Milano.
- Esser F., Umbricht A. (2013). "Competing models of journalism? Political affairs coverage in US, British, German, Swiss, French and Italian newspapers", *Journalism*, 14(8), 989-1007.
- Epstein. E.J. (1973), *News from Nowhere*, Hastings, New York.
- Fourcade M., Healy K. (2017), "Seeing like a market", *Socio-Economic Review*, 15(1), 9-29.
- Funkhouser G.R. (1973), "The Issues of the Sixties: An Exploratory Study in the Dynamics of Public Opinion", *Public Opinion Quarterly*, 50, 533-538.

- Gandy O. jr. (1998), *Communication and Race: A Cultural Perspective*, Oxford University Press, New York.
- Gerli M. (2024), *Bourdieu, Latour e la sociologia del giornalismo. Prospettive di ricerca*, Orthotes, Napoli-Salerno.
- Gerli M. (2026), Per una sociologia del campo demoscopico. Patrick Champagne e la costruzione dell'opinione pubblica, in *Fare l'opinione. Media, sondaggi e potere simbolico*, Armando, Roma (in corso di pubblicazione).
- Habermas J. (1962), *Strukturwandel der Öffentlichkeit*, H. Luchterhand Verlag, Neuwied, trad. it. *Storia e critica dell'opinione pubblica*, Laterza, Roma-Bari, 1977.
- Hallin D.C., Mancini P. (2004), *Comparing Media Systems. Three Models of Media and Politics*, Cambridge University Press, Cambridge.
- Hartsock J.C. (2000), *A History of American Literary Journalism. The Emergence of a Modern Narrative Form*, University of Massachusetts Press, Amherst.
- Janowitz M. (1975), "Professional models in journalism: The gatekeeper and the advocate", *Journalism Quarterly*, 52, 618-626.
- Jones B.D., Baumgartner, F.R. (2005), *The Politics of Attention. How Government Prioritizes Problems*, Chicago University Press, Chicago.
- Kuypers J.A. (2014), *Partisan Journalism: a History of Media Bias in United States*, Rowman & Littlefield, New York
- Lippmann W. (1922), *Public Opinion*, Macmillan, New York.
- Ludovisi D. (2016), *Il potere dei dati. Il data journalism e le nuove forme del comunicare*, Effequ, Orbetello.
- Lupton D. (2014), *Digital sociology*, Routledge, London & New York.
- Mancini P. (2002), *Il sistema fragile*, Carocci, Roma.
- Marini R. (2006), *Mass media e discussione pubblica. Le teorie dell'agenda setting*, Laterza, Roma-Bari.
- Marini R. (2014), "Per la verità: il giornalismo secondo 'Public Opinion'", *Problemi dell'informazione*, 3, 385-414.
- Marini R. (2021), "Ragionando di *advocacy* nel giornalismo. La difesa dei diritti alla prova dell'immigrazione e della società multiculturale", *Problemi dell'informazione*, 1, 187-212.
- Marini R., Gerli M. (2017), "Le forme di un tema. L'immigrazione nell'arena dei *quality papers* italiani", *Comunicazione Politica*, 3, 481-506.
- Marini R. Gerli M. (2023), "Giornalismo e immigrazione nei crocevia delle crisi", *Problemi dell'Informazione*, 3, 393-424.
- Marini R., Tusini S., Gerli M., Marcaccio A. (2019), *Immigrazione e sicurezza. Riflessioni sociologiche su politiche, rappresentazioni, linguaggi*, Maggioli, Santarcangelo di Romagna.
- Marino G.D. (1974), *La formazione dello spirito borghese in Italia*, La Nuova Italia, Firenze.
- Marradi A., Gasperoni G. (a cura di) (2002), *Costruire il dato. Vizi e virtù di alcune tecniche di raccolta delle informazioni*, FrancoAngeli, Milano.
- Marradi A. (2007), *Metodologia delle scienze sociali*, Il Mulino, Bologna.
- Mau S. (2020), "Numbers matter! The society of indicators, scores and ratings", *International Studies in Sociology of Education*, 29(1-2), 19-37.

- McLuhan M. (1962), *The Gutenberg Galaxy*, Toronto University Press, Toronto.
- Mennicken A., Espeland W.N. (2019), “What’s New with Numbers? Sociological Approaches to the Study of Quantification”, *Annual Review of Sociology*, 45(1), 223-245.
- Meyer P. (1973), *Precision Journalism: A Reporter’s Introduction to Social Science Methods*, Indiana University Press, Bloomington.
- Meyer P. (2004), *The Vanishing Newspaper: Saving Journalism in the Information Age*, University of Missouri Press, Columbia.
- Natale P. (2009), *Attenti al sondaggio*, Laterza, Roma-Bari.
- Nerone J.C. (1995), *Last Rights: Revisiting Four Theories of the Press*, University of Illinois Press, Urbana-Champaign.
- Nimmo D.D. (1970), *Political Persuaders: The Techniques of Modern Election Campaigns*, Transaction Publishers, Piscataway.
- Petre C. (2013), *A Quantitative Turn in Journalism?*, Tow Center for Digital Journalism, October 30.
- Porlezza C. (2016), “Dall’open journalism all’open government? Il ruolo del data journalism nella trasparenza e nella partecipazione”, *Problemi dell’informazione*, 1, 167-194.
- Price V. (1992), *Public Opinion*, Sage, Thousand Oaks.
- Power M. (1994), *The Audit Explosion*, Demos, London.
- Sabato L. (1981), *The Rise of Political Consultants. New Ways of Winning Elections*, Basic Books, New York.
- Sapiro G. (2003), “Forms of politicization in the French literary field”, *Theory and Society*, 32(5/6), 633-652.
- Scalea D. (2019), *Immigrazione. Le ragioni dei populist*, Historica, Firenze.
- Schudson M. (1978), *Discovering the News. A Social History of American Newspapers*, Basic Books, New York.
- Schudson M. (1995), *The Power of News*, Harvard University Press, Cambridge (MA).
- Sciortino G., Colombo A. (2004), “The flows and the flood: the public discourse on immigration in Italy, 1969-2001”, *Journal of Modern Italian Studies*, 9(1), 94-113.
- Siebert F.S., Peterson T., Schramm W. (1956), *Four Theories of the Press. The Authoritarian, Libertarian, Social Responsibility, and Soviet Communist Concepts of What the Press Should Be and Do*, University of Illinois Press, Urbana-Champaign.
- Sorrentino C. (2008), *La società densa. Riflessioni intorno alle nuove forme di sfera pubblica*, Le Lettere, Firenze.
- Sorrentino C. (2025), *Il giornalismo ha un futuro. Perché sta cambiando, come va ripensato*, Il Mulino, Bologna.
- Thompson J.B. (1995), *The Media and Modernity: a Social Theory of the Media*, Polity Press, Cambridge.
- Totolo F. (2024), *Le vite delle donne contano*, Altafortedizioni, Roma.
- Tuchman G. (1972), “Objectivity as Strategic Ritual: An Examination of News-men’s Notions of Objectivity”, *American Journal of Sociology*, (77) 4, 660-679.

- Tuchman G. (1978), *Making News: A Study in the Construction of Reality*, Free Press, New York.
- Valentini E. (2025), *Sociologia delle piattaforme digitali. Percorsi del giornalismo contemporaneo*, Carocci, Roma.
- van Dijck J., Poell T., de Waal M. (2018), *The Platform Society*, Oxford University Press, Oxford & New York.
- Watkins J.W.N. (1986), *Certezza e verità. Per una epistemologia post-popperiana*, Laterza, Roma-Bari.
- Zajczyk F. (1994), *La conoscenza sociale del territorio. Fonti e qualità dei dati*, FrancoAngeli, Milano.
- Zaller J. (1992), *The Nature and Origins of Mass Opinion*, Cambridge University Press, Cambridge.

6. *Pensare, sentire, agire l'algoritmo: un'analisi qualitativa della consapevolezza algoritmica*

di *Enrica Amaturò, Ciro Clemente De Falco, Caterina Ambrosio**

1. Introduzione

Negli ultimi anni, il dibattito sull'impatto dell'intelligenza artificiale e degli algoritmi sulla vita sociale e politica si è progressivamente spostato da approcci che enfatizzavano gli effetti unidirezionali degli algoritmi sugli individui, a una prospettiva centrata sulle condizioni di possibilità dell'agency degli utenti all'interno di ambienti digitali governati da sistemi opachi di raccomandazione, ranking e moderazione. In questo scenario, uno dei rimedi più spesso invocati è l'investimento in competenza e consapevolezza, intese non soltanto come capacità di utilizzare strumenti digitali, ma come abilità di riconoscere in che modo le mediazioni automatizzate orientino ciò che diventa visibile, rilevante, premiato o penalizzato.

Questa direzione è ormai esplicita anche a livello europeo, dove l'AI Act introduce un obbligo di «*AI literacy*» (Art. 4), imponendo a provider e deployer di garantire un livello adeguato di alfabetizzazione all'IA e agli algoritmi per gli utilizzatori in ambito organizzativo, in relazione ai contesti d'uso e ai gruppi impattati. Parallelamente, il Digital Services Act interviene sul cuore della mediazione algoritmica nelle piattaforme, richiedendo trasparenza sui principali parametri dei *recommender systems* e sulla possibilità, per gli utenti, di modificarli o influenzarli (Art. 27)¹, e-per le piattaforme molto grandi, di offrire almeno un'opzione non basata su profilazione (Art. 28). Nello stesso solco, il Digital Education Action Plan 2021-2027

* Il capitolo è il risultato della riflessione congiunta delle autrici e dell'autore.

¹ <https://dsa-observatory.eu/2024/11/22/the-regulation-of-recommender-systems-under-the-dsa-a-transition-from-default-to-multiple-and-dynamic-controls/>

inserisce la «*digital literacy*» e l'attenzione etica a dati e IA come pilastri della trasformazione educativa europea².

Se le istituzioni convergono sull'urgenza di promuovere l'alfabetizzazione all'IA e di rendere più trasparenti i sistemi di raccomandazione, la letteratura mostra tuttavia che la consapevolezza algoritmica non costituisce un oggetto semplice né unidimensionale. Più che una variabile presente/assente, essa tende a manifestarsi come un insieme eterogeneo di elementi cognitivi, interpretativi e pratici: gli individui sviluppano spesso spiegazioni di senso comune del funzionamento degli algoritmi (Ytre-Arne e Moe 2021) costruite per tentativi, scambi tra pari, osservazione di esiti e interpretazione di segnali deboli della piattaforma (Zhao, 2025).

Questa complessità si intreccia con una seconda evidenza: la consapevolezza è distribuita in modo diseguale e può riflettere linee di differenziazione che ricordano il digital divide. Inoltre, anche quando gli utenti capiscono che la visibilità è mediata da sistemi automatizzati, ciò non implica che siano disponibili, o desiderabili, strategie efficaci per intervenire su tali meccanismi. Il riconoscimento della mediazione algoritmica non si traduce automaticamente in specifiche modalità di azione, poiché le possibilità di intervento degli utenti sono strutturalmente condizionate dal regime di visibilità prodotto dalle piattaforme. Da un lato, tale regime combina *gatekeeping* e sorveglianza, limitando la leggibilità e la manipolabilità dei meccanismi algoritmici; dall'altro, esso è intrecciato a logiche economiche e politiche che incidono sulla circolazione delle narrazioni e sulla configurazione dello spazio pubblico digitale. In questo quadro, l'agenzia degli utenti assume la forma di una negoziazione situata all'interno di infrastrutture opache (Gandini *et al.*, 2025).

È proprio in questo quadro che si colloca un gap empirico ancora rilevante poiché molte ricerche, così come numerosi interventi di policy, tendono ad assumere implicitamente che un aumento della consapevolezza algoritmica produca automaticamente pratiche più autonome, critiche o democratiche. Tuttavia, le evidenze disponibili suggeriscono che il rapporto tra consapevolezza e comportamento non sia lineare né univoco. Anche in presenza di una consapevolezza diffusa della mediazione algoritmica, possono infatti emergere configurazioni differenti di relazione tra attori e sistemi algoritmici, che non possono essere dedotte a partire dal solo livello di conoscenza o informazione posseduto. In questo senso, la consapevolezza non rappresenta un punto di arrivo, ma piuttosto un campo di possibilità entro cui si articolano pratiche, interpretazioni e orientamenti normativi eterogenei.

² <https://education.ec.europa.eu/focustopics/digitaleducation/actions>

A partire da questa premessa, il presente studio si basa su un corpus di interviste in profondità condotte con un campione di attivisti e attiviste, al fine di esplorare come la mediazione algoritmica venga compresa, interpretata e, soprattutto, tradotta in pratiche comunicative. L'obiettivo non è misurare un livello di consapevolezza, ma ricostruire configurazioni ricorrenti di pratiche e cornici di senso che emergono quando l'azione degli attori si confronta con regimi di visibilità algoritmica. In questo quadro, l'attivismo costituisce un osservatorio particolarmente informativo, poiché consente di osservare in modo esplicito dinamiche di adattamento, negoziazione e resistenza alla mediazione algoritmica che, in altri contesti d'uso, rimangono spesso meno visibili o meno tematizzate.

A tal fine, il capitolo è organizzato come segue. Nel secondo paragrafo viene proposto un inquadramento teorico della consapevolezza algoritmica, mettendone in evidenza la natura multidimensionale e i principali aspetti che, alla luce della letteratura esistente, meritano ulteriore indagine empirica. Il terzo paragrafo descrive le caratteristiche metodologiche della ricerca e il disegno empirico adottato. Nel quarto paragrafo sono presentati e discussi i risultati dell'analisi, mentre il paragrafo conclusivo è dedicato alla riflessione sulle principali implicazioni teoriche e interpretative di quanto emerso.

2. Studiare la consapevolezza algoritmica

2.1. Modelli concettuali della consapevolezza algoritmica

Le modalità attraverso cui la consapevolezza algoritmica viene studiata in letteratura sono molteplici, dando luogo a una notevole eterogeneità di approcci che rende complesso il confronto sistematico tra i diversi contributi (Felaco, 2024). Tale pluralità emerge già a partire dalla terminologia impiegata per descrivere il fenomeno, che riflette definizioni, sfumature concettuali e dimensioni analitiche differenti a seconda dei contesti e degli obiettivi di ricerca adottati.

Se si considerano, ad esempio, le etichette utilizzate per definire questo fenomeno, si osserva una coesistenza di diverse espressioni, tra le quali le più ricorrenti sono consapevolezza algoritmica e *algorithmic literacy*. Sebbene il termine consapevolezza venga spesso impiegato per riferirsi prevalentemente ad aspetti legati alla percezione degli algoritmi (Dogruel *et al.*, 2021) o alla comprensione del funzionamento dei meccanismi algoritmici (Alvarado, 2019), in questo lavoro si adotta l'espressione consapevolezza algoritmica (CA). Riprendendo quanto scritto da Felaco (2022), si condivide l'idea per cui l'espressione «consapevolezza algoritmica», delinea il fenomeno in senso

ampio, e quella di «alfabetizzazione algoritmica» le azioni e agli interventi orientati all'educazione e alla sensibilizzazione dei cittadini.

Oltre alla variabilità terminologica, anche le teorizzazioni della consapevolezza algoritmica presentano una notevole eterogeneità, venendo spesso articolate attraverso modelli multidimensionali che riflettono differenti assunti teorici e obiettivi analitici. Nel lavoro dedicato allo sviluppo e alla validazione di una scala di alfabetizzazione algoritmica, Dogruel *et al.* (2022) propongono una concettualizzazione di tipo multilivello. A seguito di una revisione della letteratura, gli autori individuano quattro dimensioni principali che definiscono l'alfabetizzazione algoritmica: (1) consapevolezza e conoscenza, intese come la consapevolezza dell'esistenza degli algoritmi e la comprensione del loro funzionamento; (2) valutazione critica, riferita alla capacità di valutare in modo critico i processi decisionali algoritmici; (3) strategie di coping, intese come strategie comportamentali adottate per gestire o contrastare gli effetti degli algoritmi, soprattutto quando questi sono percepiti come fonte di pressione, incertezza o sovraccarico informativo (Kamble *et al.*, 2026); e (4) competenze di creazione e progettazione, che includono le capacità di modificare o influenzare direttamente il funzionamento dei sistemi algoritmici. Tuttavia, nella fase empirica dello studio, gli autori considerano esclusivamente le prime due dimensioni, poiché ritengono che le restanti richiederebbero strumenti di misurazione più complessi. In modo analogo, Zarouali *et al.* (2021), a partire da una revisione della letteratura sugli algoritmi di raccomandazione, personalizzazione e filtraggio dei contenuti, definiscono nella fase teorica cinque dimensioni costitutive della *algorithmic media content awareness*: (1) consapevolezza del filtraggio dei contenuti; (2) consapevolezza delle decisioni automatizzate; (3) consapevolezza dell'interazione uomo-algoritmo; (4) consapevolezza della persuasione algoritmica; e (5) consapevolezza delle implicazioni etiche. Tuttavia, attraverso analisi fattoriali esplorative e confermate condotte su campioni rappresentativi di utenti, il modello empirico finale si stabilizza attorno alle prime quattro dimensioni. Felaco (2022) propone invece un modello teorico a quattro dimensioni, volto a sistematizzare e rendere più operativa la nozione di consapevolezza algoritmica. Sulla base di una revisione dei principali contributi internazionali (tra cui Cotter e Reisdorf, 2020; Zarouali *et al.*, 2021; Dogruel *et al.*, 2022), l'autore concepisce la consapevolezza algoritmica come un costrutto articolato in quattro dimensioni interrelate: cognitiva, comportamentale, procedurale e affettiva. La dimensione cognitiva riguarda i modi in cui gli utenti diventano consapevoli dell'esistenza e del funzionamento degli algoritmi, includendo sia una consapevolezza generale sia una comprensione più approfondita delle logiche che li governano. La dimensione comportamentale fa invece riferi-

mento alle strategie e alle tattiche che gli utenti adottano per adattarsi alle scelte algoritmiche o per reagire ad esse. La dimensione procedurale implica la capacità di valutare criticamente il funzionamento degli algoritmi, riconoscendone i rischi e le implicazioni etiche, mentre la dimensione affettiva comprende le reazioni emotive suscitate dall'interazione con sistemi automatizzati. Questo modello teorico viene concepito dall'autore come una cornice concettuale flessibile, pensata per orientare futuri tentativi di analisi empirica della consapevolezza algoritmica.

Questi rappresentano solo alcuni esempi delle diverse concettualizzazioni adottate in letteratura. Non esiste, ovviamente, un approccio corretto o errato allo studio della consapevolezza algoritmica, poiché le scelte teoriche e metodologiche dipendono in larga misura dagli obiettivi di ricerca e dall'impostazione analitica adottata.

Il quadro delle diverse concettualizzazioni risulta ulteriormente complesso se messo in relazione con quanto discusso nell'introduzione. Il rapporto tra le varie dimensioni della consapevolezza algoritmica non è lineare, né il suo accrescimento segue un percorso cumulativo e progressivo. In questa prospettiva, la consapevolezza algoritmica non può essere intesa unicamente come un obiettivo da raggiungere, ma piuttosto come un campo di possibilità entro cui si articolano pratiche, interpretazioni e orientamenti normativi eterogenei. È proprio per questo motivo che la dimensione comportamentale e quella valutativa risultano particolarmente in linea con gli obiettivi del lavoro.

2.2. Approcci empirici e modalità di restituzione dei risultati

La consapevolezza algoritmica viene studiata con approcci sia qualitativi che quantitativi. Una modalità particolarmente diffusa tra gli approcci quantitativi consiste nell'impiego di strumenti di tipo *self-report*, in cui agli individui viene richiesto di auto-posizionarsi lungo una scala di consapevolezza. Questo approccio è adottato, ad esempio, nello studio di Gran *et al.* (2021). Ai partecipanti di questo lavoro è stato chiesto di dichiarare il proprio grado di conoscenza rispetto a diversi specifici aspetti dei meccanismi algoritmici. Tuttavia, diversi autori hanno evidenziato i limiti di questa modalità di rilevazione, riconducibili principalmente al bias di desiderabilità sociale, che può compromettere l'affidabilità delle risposte e condurre a una sovrastima dei livelli di consapevolezza auto-riferiti (Felaco e Pelliccia, 2024). Alla luce di tali criticità, alcuni contributi hanno scelto di ricorrere al questionario adottando approcci alternativi al *self-report*. Questo avviene, ad esempio, nello studio di Dogruel *et al.* (2022). Gli autori partono dall'i-

dea che chiedere agli individui quanto si sentano consapevoli degli algoritmi rischi di produrre misure distorte, fortemente influenzate da aspetti come la sovra- o sottovalutazione soggettiva e la desiderabilità sociale. Per questo motivo, il fenomeno viene trattato prevalentemente sotto l'aspetto cognitivo, osservabile attraverso prestazioni e risposte verificabili, piuttosto che come una percezione dichiarata. Il questionario sviluppato dagli autori si basa infatti su item, formulati come domande vero/falso o a scelta multipla, essendo, di fatto un test per la verifica delle competenze.

Le interviste rappresentano la tecnica maggiormente adottata tra gli approcci qualitativi. Uno dei contributi più ampi in questo ambito è quello di Hargittai *et al.* (2020). Sono state condotte oltre 80 interviste individuali in cinque paesi diversi. Tali interviste sono state arricchite con esercizi di disegno e discussione, al fine di esplorare soprattutto le modalità con cui gli individui interpretano il funzionamento algoritmico. Un altro esempio in questa direzione è rappresentato dallo studio di Swart (2021). Questo lavoro adotta un disegno qualitativo esplorativo basato su interviste in profondità con 22 giovani. Le interviste sono integrate con il *walk-through method* e con *think-aloud protocols*. Questi approcci prevedono la navigazione guidata delle piattaforme social mentre i partecipanti verbalizzano in tempo reale le proprie interpretazioni. Tra gli altri strumenti qualitativi, un filone particolarmente rilevante è quello di matrice etnografica. Un esempio è fornito da Pronzato e Markham (2023), che hanno coinvolto 40 studenti in un processo di autoetnografia guidata, finalizzato a esplorare le modalità attraverso cui gli individui sperimentano e interpretano la presenza degli algoritmi nei media digitali all'interno della vita quotidiana. Lo studio si fonda sull'uso di diari auto etnografici strutturati, integrati con esercizi di scrittura riflessiva, tecniche di *visual elicitation* e pratiche di disconnessione temporanea, configurandosi come un dispositivo metodologico orientato a stimolare forme di consapevolezza algoritmica di tipo critico.

Anche nel lavoro di Siles *et al.* (2022), la consapevolezza algoritmica viene indagata attraverso l'impiego di diari etnografici, utilizzati per ricostruire il modo in cui gli utenti sviluppano nel tempo una relazione con gli algoritmi di TikTok. Gli autori hanno coinvolto 43 nuovi utenti della piattaforma, chiedendo loro di compilare diari di utilizzo per un periodo di un mese, nei quali annotare pratiche, percezioni, emozioni e interpretazioni legate alle raccomandazioni algoritmiche. A questi materiali diaristici sono stati affiancati sette focus group, svolti sia prima sia dopo la fase dei diari, con una funzione di contestualizzazione e triangolazione dei dati. Tale disegno multi-tecnica consente di analizzare la consapevolezza algoritmica non come una proprietà statica, ma come un processo dinamico che si articola nel tempo lungo dimensioni cognitive, affettive e pratiche.

Sono relativamente pochi i contributi che propongono una restituzione dei risultati sotto forma di tipi qualitativamente caratterizzati. Uno dei rari esempi in questa direzione è rappresentato dalla classificazione delle *folk-theories*³ dei partecipanti al suo studio proposta da DeVito (2021). L'autrice propone una classificazione ottenuta attraverso un'analisi di tipo *grounded* dei materiali raccolti. I diversi tipi che l'autrice individua spaziano dalla *basic awareness* fino a forme più articolate di *mechanistic ordering*, descritte attraverso interpretazioni e spiegazioni fornite dagli utenti. In questo caso, la consapevolezza algoritmica non viene trattata come una variabile metrica, ma come una tipologia qualitativa emergente dalle pratiche di *sensemaking* degli utenti.

Come emerge dai contributi discussi, la scelta delle tecniche di raccolta dei dati dipende in modo sostanziale dagli obiettivi della ricerca. Nel caso dei questionari, l'obiettivo principale è spesso quello di ottenere risultati standardizzati e comparabili, generalmente analizzati in relazione a variabili sociodemografiche. Al contrario, gli approcci qualitativi mirano a comprendere come le persone percepiscono, interpretano e negoziano la presenza degli algoritmi nella loro vita quotidiana, incoraggiando al contempo una riflessione critica sui processi di *datafication* e sorveglianza e ponendo particolare attenzione alla dimensione temporale (idem). In virtù della complementarità degli aspetti colti da questi diversi approcci, diversi autori suggeriscono di combinare tecniche quantitative e qualitative all'interno di disegni di ricerca di tipo *mixed methods*, al fine di integrare la profondità interpretativa con il potenziale di generalizzazione dei risultati (Dogruel *et al.*, 2022; Lupton e Michael, 2017).

In generale è possibile affermare che in una parte consistente della letteratura empirica, la consapevolezza algoritmica viene trattata come una proprietà metrica, misurabile lungo un continuum quantitativo e restituita sotto forma di livelli o punteggi. Gli studi di natura qualitativa evitano questo approccio e tendono invece a restituire i risultati sotto forma di "temi emergenti". Un esempio in tal senso è offerto dallo studio di Siles *et al.* (2022). La consapevolezza algoritmica viene analizzata come un processo e restituita al lettore attraverso delle attività ricorrenti che caratterizzano il comportamento dei partecipanti alla ricerca.

Alla luce di questo quadro, in questo lavoro si propone non di misurare

³ Le *folk theories* sono le concezioni che le persone utilizzano nella vita quotidiana per dare senso al funzionamento del mondo sociale. Nella fattispecie questo approccio non si concentra su come le persone usano effettivamente i media o interpretano messaggi specifici, ma sulle idee e sulle credenze che esse hanno riguardo a come i media funzionano, ovvero sulle loro "teorie" (Ytre-Arne e Moe, 2021).

un livello di consapevolezza, ma ricostruire configurazioni ricorrenti di pratiche e cornici di senso, indagando quindi in modo specifico la dimensione comportamentale e quella valutativa. La prima è intesa in termini di azioni concrete e forme di agency che gli individui esercitano nell'interazione con i sistemi algoritmici, mentre la seconda riguarda le valutazioni attribuite agli algoritmi, sia sul piano dell'esperienza personale sia in relazione ai loro impatti sociali più ampi. In virtù della natura sfumata e processuale di tali dimensioni, si ritiene che la restituzione dei risultati attraverso una tipologia qualitativa rappresenti l'approccio più adeguato, che per l'appunto manca in letteratura. L'obiettivo è quello di cogliere la complessità delle pratiche, delle interpretazioni e delle valutazioni degli individui, offrendo una rappresentazione articolata della consapevolezza algoritmica che difficilmente potrebbe essere ricondotta a punteggi o livelli standardizzati.

3. Disegno della ricerca e strategia di campionamento

Alla luce del quadro teorico delineato, la ricerca adotta un disegno qualitativo esplorativo, ritenuto particolarmente adeguato a indagare dimensioni della consapevolezza algoritmica difficilmente riconducibili a un approccio di tipo metrico. In linea con l'obiettivo di ricostruire pratiche, interpretazioni e giudizi attribuiti agli algoritmi, lo studio si basa su un corpus di interviste semi-strutturate in profondità, concepite come strumento privilegiato per cogliere processi, riflessività e modalità situate di relazione con i sistemi algoritmici.

Considerate le domande di ricerca, si è ritenuto necessario selezionare un campione di soggetti che, per il tipo di attività svolta, utilizzassero in modo sistematico i social media e che, proprio in virtù di tale utilizzo, avessero con buona probabilità sviluppato una riflessione sul funzionamento degli algoritmi, sia in termini propositivi (strategie di visibilità, diffusione dei contenuti) sia reattivi (adattamento, critica, resistenza). In altre parole, l'interesse non era rivolto a utenti medi delle piattaforme, ma a soggetti che, per posizione e pratiche, fossero esposti in modo più intenso e consapevole alla mediazione algoritmica.

In continuità con precedenti studi di caso che hanno indagato la consapevolezza algoritmica in contesti caratterizzati da un uso intensivo e strategico delle piattaforme digitali (ad esempio ricerche focalizzate su influencer o creator (Bonini e Treré, 2024), si è scelto di concentrare l'attenzione su attivisti politici che integrano l'uso dei social media nelle proprie pratiche di azione pubblica. L'attivismo rappresenta infatti un osservatorio particolarmente informativo per l'analisi della consapevolezza algoritmica,

poiché combina un'esposizione diretta ai regimi di visibilità delle piattaforme con una marcata riflessività sulle implicazioni politiche, etiche e sociali degli strumenti utilizzati. In questo senso, la selezione del campione risponde a una logica teorica piuttosto che rappresentativa, poiché l'obiettivo non è generalizzare i risultati a una popolazione più ampia, ma esplorare in profondità configurazioni di pratiche e cornici di senso in un contesto ad alta densità di interazione con gli algoritmi.

La selezione dei partecipanti è avvenuta attraverso una strategia di campionamento intenzionale (*purposive sampling*), combinata con una successiva logica a valanga (*snowball sampling*). In una prima fase sono stati individuati alcuni contatti primari tramite le reti personali dei ricercatori e mediante il contatto diretto con potenziali partecipanti attraverso i canali social. A partire da questi primi interlocutori, sono stati poi raccolti ulteriori contatti seguendo una logica a valanga, che ha permesso di ampliare il campione includendo soggetti appartenenti a contesti di attivismo differenti ma accomunati dall'uso dei social media come strumento di comunicazione e mobilitazione politica.

Le interviste sono state condotte nella provincia di Napoli nel periodo aprile-maggio 2024 e hanno coinvolto complessivamente 20 attivisti e attiviste. Le interviste, della durata compresa tra un minimo di circa 30 minuti e un massimo di un'ora, sono state realizzate in forma semi-strutturata, seguendo una traccia che toccava esplicitamente le dimensioni emerse dalla letteratura sulla consapevolezza algoritmica (conoscenza, pratiche, valutazioni e percezioni degli effetti sociali degli algoritmi). Tutte le interviste sono state registrate previo consenso informato dei partecipanti e successivamente trascritte per l'analisi. L'obiettivo dell'analisi non è quello di misurare o graduare la consapevolezza algoritmica, ma di identificare configurazioni ricorrenti di pratiche e cornici di senso che emergono nell'interazione tra attori e sistemi algoritmici. A tal fine, il materiale empirico è stato analizzato con l'intento di costruire una tipologia qualitativa di approcci alla mediazione algoritmica, ottenuta attraverso l'incrocio di dimensioni analitiche teoricamente informate ed empiricamente articolate.

4. Dimensioni analitiche e costruzione della tipologia

4.1. Articolazione empirica delle dimensioni analitiche

Questa sezione presenta le dimensioni analitiche alla base della costruzione della tipologia degli approcci degli attivisti ai sistemi algoritmici. L'obiettivo dell'analisi non è quello di classificare gli individui in profili

rigidi, ma di individuare tipi analitici (configurazioni ricorrenti di pratiche e cornici interpretative) attraverso cui gli attivisti interagiscono con, attribuiscono significato a e rispondono alla mediazione algoritmica nell'azione politica.

Le dimensioni analitiche utilizzate per la costruzione della tipologia sono due: ingaggio algoritmico e cornici di senso dell'algoritmo. Con ingaggio algoritmico si intendono le modalità attraverso cui gli attori sociali rispondono – o non rispondono – ai sistemi algoritmici delle piattaforme digitali al fine di orientare le proprie pratiche comunicative. Con cornici di senso dell'algoritmo si intendono gli schemi attraverso cui gli attori sociali interpretano e valutano i sistemi algoritmici delle piattaforme digitali.

Queste non emergono in modo puramente induttivo ma risultano incorporate nella traccia d'intervista, progettata per sollecitare esperienze, pratiche e interpretazioni legate al ruolo degli algoritmi nella comunicazione e nella mobilitazione politica. Ciò che emerge dall'analisi qualitativa sono invece le categorie specifiche assunte da ciascuna dimensione e, soprattutto, i tipi analitici che derivano dal loro intreccio. Nella tabella 1 sono sintetizzati i risultati dell'analisi qualitativa delle interviste, articolati in dimensioni analitiche e categorie interpretative emerse dall'esame del materiale empirico. Per ciascuna categoria è stato selezionato, a titolo esemplificativo, uno stralcio di intervista al fine di rendere esplicito il legame tra le categorie analitiche e le narrazioni degli intervistati.

Tab. 1 - Dimensioni analitiche e categorie emerse dall'analisi qualitativa

Dimensione	Categoria	Citazione
Ingaggio algoritmico	<i>Adattamento</i>	«Postiamo ogni due giorni per far prendere aria al post. Usiamo gli audio di tendenza nei video, magari li mettiamo a volume basso, però servono a far girare il contenuto.» (Int. 2, attivismo studentesco)
	<i>Elusione</i>	«Per far passare contenuti sulla Palestina abbiamo iniziato a cambiare le parole, a scriverle in modo diverso, oppure a nasconderle dentro immagini o video che sembravano altro.» (Int. 1, collettivo politico)
	<i>Non-compliance selettiva (refusal)</i>	«Sappiamo cosa potrebbe andare virale e cosa no, però non vogliamo abbassarci più di tanto. Preferiamo subire le conseguenze senza distorcere contenuti e messaggi.» (Int. 7, attivismo queer)
	<i>Indifferenza</i>	«La militanza vera non si fa sui social. I social possono servire, ma se non attraversi gli spazi, se non stai con la gente, non stai facendo politica.» (Int. 9, attivismo territoriale)

Cornici di senso dell'algoritmo	<i>Pragmatica</i>	«Riconosco che la diffusione avviene se stai alle regole del gioco. O lo dirigi un po', oppure lo subisci.» (Int. 5, attivismo civico)
	<i>Ambivalente</i>	«Gli algoritmi possono essere utili per la visibilità, però allo stesso tempo ti costringono a essere più performativo, a fare cose che poi politicamente non servono a niente.» (Int. 2, attivismo studentesco)
	<i>Critica radicale</i>	«Le piattaforme non sono riformabili: sono costruite per normalizzare e neutralizzare il conflitto.» (Int. 16, collettivo politico)

Come è possibile notare dalla tabella, relativamente all'ingaggio algoritmico, quattro sono le modalità specifiche di ingaggio emerse dall'analisi delle interviste.

- a) *L'adattamento* comprende le pratiche orientate all'ottimizzazione della visibilità e della circolazione dei contenuti politici operando all'interno delle regole implicite delle piattaforme. Rientrano in questa modalità l'attenzione ai formati, ai tempi di pubblicazione, alla frequenza dei post e all'adozione di stili o trend percepiti come algoritmicamente premianti.
- b) *L'elusione* comprende le pratiche attraverso cui gli attivisti cercano di aggirare meccanismi di moderazione, filtraggio o penalizzazione della visibilità associati a temi politicamente sensibili. Queste pratiche includono la modifica intenzionale di parole chiave o hashtag, l'uso di linguaggi codificati o simbolici e l'inserimento di messaggi politici all'interno di formati apparentemente neutri.
- c) *La non-compliance selettiva* indica una scelta consapevole di non adottare strategie di ottimizzazione o di elusione, pur essendone consapevoli dell'efficacia potenziale. In questo caso, gli attivisti accettano deliberatamente una riduzione della visibilità per preservare la coerenza politica, l'integrità del linguaggio o specifici principi etici.
- d) Infine, *l'indifferenza* descrive situazioni in cui gli algoritmi non assumono un ruolo rilevante nell'orientare le pratiche politiche. In questi casi non emergono strategie di adattamento né di elusione e l'algoritmo rimane sullo sfondo, a favore di forme di azione prevalentemente offline o relazionali.

Con riferimento alle cornici di senso dell'algoritmo, sono stati individuate tre principali cornici di senso.

- a) Una *cornice pragmatica*, in cui l'algoritmo è concepito prevalentemente come uno strumento inevitabile o come un vincolo tecnico con cui è necessario confrontarsi, senza essere tematizzato come attore politico a pieno titolo.
- b) Una *cornice ambivalente*, caratterizzata dal riconoscimento esplicito

dei trade-off tra visibilità e coerenza politica. Gli algoritmi sono visti come strumenti abilitanti e, allo stesso tempo, come dispositivi potenzialmente depolitizzanti.

- c) Infine, una *cornice critica radicale*, in cui gli algoritmi sono interpretati come infrastrutture di potere strutturalmente incompatibili con una politica emancipativa, associate a logiche di controllo, estrazione e normalizzazione. I riferimenti agli effetti sociali e politici degli algoritmi – come censura, silenziamento o distorsione del dibattito pubblico – sono trattati come elementi costitutivi di queste cornici di senso, e non come esiti oggettivamente misurati dei sistemi algoritmici.

Le dimensioni e le categorie fin qui discusse consentono di cogliere separatamente le modalità pratiche di ingaggio algoritmico e le cornici interpretative attraverso cui gli attivisti attribuiscono significato politico agli algoritmi. Tuttavia, tali elementi non si presentano empiricamente come dimensioni isolate, ma tendono a combinarsi in modo ricorrente nelle pratiche e nelle narrazioni degli intervistati. L'analisi congiunta delle due dimensioni permette quindi di identificare configurazioni ricorrenti, che vengono qui ricostruite sotto forma di tipi analitici. Nel paragrafo che segue vengono presentati i quattro tipi individuati, che sintetizzano le principali modalità attraverso cui gli attivisti si rapportano alla mediazione algoritmica nell'azione politica. Trasversalmente a tutti i tipi analitici emerge, seppur declinata in modo differente, una tensione ricorrente tra visibilità e coerenza politica, che costituisce uno degli assi principali lungo cui gli attivisti negoziano il rapporto con la mediazione algoritmica.

4.2. Dalle categorie ai tipi analitici

Nella tabella 2 sono sintetizzati i risultati dell'analisi tipologica, ottenuti dall'incrocio delle due dimensioni analitiche principali – l'ingaggio algoritmico e le cornici di senso dell'algoritmo. La tabella presenta i quattro tipi analitici individuati, mettendo in relazione per ciascun tipo la modalità prevalente di ingaggio algoritmico, la cornice di senso dominante e la logica d'azione che ne deriva.

Tab. 2 - *Tipi analitici emersi dall'analisi tipologica*

Tipo analitico	Ingaggio algoritmico prevalente	Cornice di senso dell'algoritmo	Logica d'azione dominante
<i>Adattivo pragmatico</i>	Adattamento	Pragmatica	Massimizzare la visibilità operando entro i vincoli della piattaforma
<i>Adattivo critico</i>	Adattamento / Elusione	Ambivalente	Negoziare visibilità e coerenza politica
<i>Resistente riflessivo</i>	Non-compliance selettiva (refusal)	Critica (non radicale)	Preservare integrità e coerenza etico-politica
<i>Rifiutante radicale</i>	Indifferenza / Rifiuto	Critica radicale	Sottrarsi all'arena algoritmica come spazio politico

Il primo tipo, definito *adattivo pragmatico*, è caratterizzato da una relazione prevalentemente strumentale con gli algoritmi, concepiti come vincoli tecnici da comprendere e gestire. In questa configurazione, l'attenzione è rivolta soprattutto all'efficacia comunicativa e alla massimizzazione della visibilità dei contenuti, mentre gli effetti politici più ampi della mediazione algoritmica rimangono sullo sfondo. L'adattamento alle logiche della piattaforma è vissuto come una competenza necessaria, raramente tematizzata in termini critici o problematizzata sul piano politico. Nel complesso, questo tipo mostra come la consapevolezza algoritmica possa tradursi in una relazione prevalentemente strumentale con le piattaforme, in cui la riflessività sugli effetti politici della mediazione rimane sullo sfondo rispetto all'obiettivo della visibilità.

Una seconda configurazione, definita di tipo *adattivo critico*, condivide l'orientamento all'adattamento ma se ne distingue per la presenza di una cornice di senso ambivalente. In questo caso, gli algoritmi sono riconosciuti come risorse indispensabili per la visibilità, ma al tempo stesso come dispositivi che impongono compromessi problematici sul piano politico. L'ingaggio algoritmico assume così la forma di una negoziazione continua tra l'esigenza di raggiungere un pubblico ampio e il tentativo di preservare la coerenza e la densità del messaggio politico. A differenza del tipo adattivo pragmatico, qui l'adattamento non è mai del tutto naturalizzato, ma rimane oggetto di riflessione critica e di tensione persistente.

Il terzo tipo, definito *resistente riflessivo*, si caratterizza invece per una presa di distanza selettiva dalle logiche algoritmiche. In questa configurazione, gli attori riconoscono gli effetti potenzialmente depoliticizzanti dell'ottimizzazione algoritmica e scelgono consapevolmente di limitarne l'adozione, assumendo la coerenza etico-politica come criterio prioritario dell'azione. La

resistenza non assume tuttavia la forma di un rifiuto totale delle piattaforme, ma si esprime attraverso pratiche di non-compliance mirate e una riflessività costante sulle conseguenze dell'ingaggio digitale. Mentre nel tipo *adattivo critico* la consapevolezza degli effetti problematici degli algoritmi non impedisce il ricorso a pratiche di adattamento ed elusione, nel tipo *resistente riflessivo* tale consapevolezza si traduce in una scelta normativa che privilegia l'integrità del messaggio rispetto all'efficacia comunicativa. A differenza dei tipi adattivi, in questa configurazione la conoscenza delle logiche algoritmiche non viene mobilitata per ottimizzare la visibilità, ma per giustificare una presa di distanza selettiva dalle pratiche ritenute politicamente problematiche.

Infine, il tipo *rifiutante radicale* si distingue per una postura di sostanziale estraneità o rifiuto dell'ingaggio algoritmico. Gli algoritmi e le piattaforme digitali sono interpretati come infrastrutture di potere strutturalmente incompatibili con una politica trasformativa, e l'azione politica è orientata prevalentemente verso spazi e pratiche esterne all'arena digitale. In questa configurazione, la mediazione algoritmica non è considerata un terreno di negoziazione, ma un ambiente da cui sottrarsi. A differenza del tipo *resistente riflessivo*, che mantiene una relazione problematica ma non interrotta con le piattaforme digitali, il tipo *rifiutante radicale* concepisce l'arena algoritmica come intrinsecamente depoliticizzante, optando per una sottrazione sostanziale dallo spazio digitale come ambito di azione politica

I quattro tipi analitici individuati mettono in relazione pratiche di ingaggio e cornici di senso, offrendo una chiave interpretativa per comprendere come la mediazione algoritmica venga riconosciuta, problematizzata e tradotta in pratiche differenti. Nel loro insieme, essi mostrano come la consapevolezza della mediazione algoritmica sia ampiamente condivisa tra gli attori, ma non sia di per sé sufficiente a spiegare o predire i comportamenti osservati.

I risultati suggeriscono infatti che la relazione tra consapevolezza, adattamento e resistenza non segua una traiettoria lineare, ma assuma la forma di configurazioni situate, che dipendono dal modo in cui gli algoritmi vengono interpretati come risorsa, vincolo o infrastruttura di potere. In questo senso, i tipi analitici non rappresentano stadi evolutivi né livelli crescenti di consapevolezza, ma modalità differenti di articolazione tra pratiche, interpretazioni e criteri normativi.

5. Riflessioni conclusive

Questo lavoro ha mostrato come la consapevolezza algoritmica si configura come un fenomeno intrinsecamente relazionale, situato e processuale, che prende forma nell'intreccio tra pratiche di ingaggio, interpretazioni e

logiche d'azione. A partire da una breve ricognizione della letteratura, è emersa con chiarezza la forte eterogeneità che caratterizza sia le definizioni sia le strategie empiriche e come spesso venga trattata come una proprietà puramente metrica e restituita attraverso punteggi, livelli o profili standardizzati.

Questo studio ha esplorato le dimensioni della consapevolezza algoritmica che risultano meno facilmente riconducibili a una logica quantitativa, ovvero la dimensione comportamentale e affettiva. In particolare, attraverso un'analisi qualitativa di tipo abducente su 20 attivisti politici presenti digitalmente, il contributo principale risiede nella costruzione di una tipologia analitica che mette in relazione l'ingaggio algoritmico e le cornici di senso. Sono individuati quattro tipi che non vanno intesi come categorie rigide né come profili esaustivi degli attori, ma come strumenti concettuali.

Il tipo *adattivo pragmatico* è caratterizzato da pratiche di ottimizzazione della visibilità e da una cornice di senso che interpreta l'algoritmo come un vincolo tecnico inevitabile, da comprendere e utilizzare in modo strumentale. Il tipo *adattivo critico* combina invece pratiche di adattamento ed elusione con una cornice ambivalente, in cui gli algoritmi sono riconosciuti simultaneamente come risorse indispensabili per la visibilità e come dispositivi potenzialmente depolitizzanti. Il tipo *resistente riflessivo* si caratterizza per una postura di non-compliance selettiva, sostenuta da una cornice critica che non implica un rifiuto totale delle piattaforme. Infine, il tipo *ri-fiutante radicale* combina una cornice di senso critica radicale con un disingaggio sostanziale dagli algoritmi e dalle piattaforme digitali, interpretate come infrastrutture di potere strutturalmente incompatibili con una politica trasformativa.

La tipologia proposta consente di mettere in luce alcune implicazioni rilevanti per lo studio della consapevolezza e della resistenza algoritmica nell'azione politica. In primo luogo, i quattro tipi analitici mostrano come il rapporto degli attivisti con i sistemi algoritmici non possa essere ridotto a una dicotomia tra adattamento e rifiuto, né interpretato esclusivamente in termini di livelli crescenti di competenza o consapevolezza. Al contrario, le configurazioni individuate evidenziano che pratiche simili possono essere sostenute da cornici di senso differenti, e che giudizi critici sugli algoritmi non si traducono necessariamente in pratiche di disingaggio o opposizione. In questa prospettiva, la distinzione tra tipi adattivi e resistenti non coincide con una distinzione tra attori consapevoli e non consapevoli, ma rimanda piuttosto a differenti modalità di negoziazione tra visibilità, coerenza politica e condizioni materiali dell'azione collettiva. Un secondo elemento rilevante riguarda il fatto che la consapevolezza algoritmica emerge come un fenomeno situato, relazionale e che varia nel tempo. I tipi individuati non

descrivono soggetti coerenti e immutabili, ma configurazioni che cambiamo forma in base a molteplici elementi. In questo senso, la tipologia consente di cogliere la natura processuale e negoziata dei fenomeni digitali. Infine, il lavoro suggerisce che approcci esclusivamente metrici rischiano di appiattire la complessità delle pratiche e delle valutazioni rispetto alla consapevolezza algoritmica.

Il lavoro non è privo di limiti. Un primo limite riguarda la dimensione temporale. Sebbene l'analisi metta in luce la natura processuale della consapevolezza algoritmica, il disegno adottato restituisce una fotografia prevalentemente sincronica delle pratiche e delle cornici interpretative degli attori. Studi futuri potrebbero adottare prospettive longitudinali per osservare come tali configurazioni evolvano nel tempo, in relazione a cambiamenti nelle infrastrutture di piattaforma, nelle politiche di moderazione o nei contesti politici più ampi. Un secondo limite concerne la dimensione spaziale della ricerca. Il lavoro si concentra infatti su un contesto geografico specifico. Tra gli altri spunti per ricerche future, si potrebbe immaginare la proposta di una tipologia in contesti differenti, sia in termini di soggetti intervistati che in termini di contesti geografici. Inoltre, come suggerito nei paragrafi precedenti, un'estensione promettente potrebbe riguardare l'integrazione tra dati qualitativi e dati quantitativi.

Nel complesso, l'analisi proposta invita a pensare la consapevolezza algoritmica come un fenomeno dinamico che si costruisce nell'interazione tra soggetti, infrastrutture digitali e contesti. In questa prospettiva, studiare la consapevolezza algoritmica significa interrogarsi non solo su ciò che gli attori sanno degli algoritmi, ma su come tale sapere venga negoziato o rifiutato e su quali sono le pratiche e le cornici di senso assegnate alle azioni.

Riferimenti bibliografici

- Alvarado O. (2019), *Breaking the Fourth Wall: Embodied Interfaces for a Better Algorithmic Experience with Recommender Algorithms*. In *Proceedings of the Thirteenth International Conference on Tangible, Embedded, and Embodied Interaction* (pp. 711-714).
- Bonini T., Treré E. (2024), *Algorithms of resistance: The everyday fight against platform power*, Mit Press, Cambridge.
- Bucher T. (2017), "The Algorithmic Imaginary: Exploring the Ordinary Affects of Facebook Algorithms", *Information, Communication & Society*, 20(1), 30-44.
- Cotter K., Reisdorf B.C. (2020), "Algorithmic Knowledge Gaps: A New Horizon of (digital) Inequality" *International Journal of Communication*, 14(21), 745-756.
- DeVito M.A. (2021), *Adaptive folk theorization as a path to algorithmic literacy on*

- changing platforms. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2), 1-38.
- DeVito M.A., Gergle D., Birnholtz J. (2017), “‘Algorithms ruin everything’ #RIPTwitter, Folk Theories, and Resistance to Algorithmic Change in Social Media”, in *Proceedings of the 2017 CHI conference on human factors in computing systems* (pp. 3163-3174).
- Dogrueel L., Facciorusso D., Stark B. (2022), “I’m Still the Master of the Machine.’Internet Users’ Awareness of Algorithmic Decision-Making and their Perception of its Effect on their Autonomy”, in *Information, Communication & Society*, 1311-1332.
- Dogrueel L., Masur P., Joeckel S. (2021), “Development and Validation of an Algorithm Literacy Scale for Internet Users”, in *Communication Methods and Measures*, 16(2), 115-133.
- Eslami M. et al. (2016), “I always assumed that I wasn’t really that close to her”: Reasoning about invisible algorithms in news feeds. *Proceedings of CHI*.
- Eslami M., Rickman A., Vaccaro K., Aleyasen A., Vuong A., Karahalios K., Sandvig C. (2015), “I always assumed that I wasn’t really that close to [her]” Reasoning about Invisible Algorithms in News Feeds”, in *Proceedings of the 33rd annual ACM conference on human factors in computing systems* (pp. 153-162).
- Felaco C. (2022), “Lungo la scala di generalità: le dimensioni della consapevolezza algoritmica”, *Sociologia Italiana*, 19/20, 123-134
- Felaco C. (2024). Researching algorithm awareness: Methodological approaches to investigate how people perceive, know, and interact with algorithms. *Mathematical Population Studies*, 31(4), 267-288.
- Felaco C., Pelliccia F. (2024), “TikTok Algorithm Awareness. The Case of Algospeak to Bypass Algorithmic Logic”, in Maiello, C.S.G., Masullo G., *The Fields of Digital Research: Teoretical, Methodological and Application Challenges* (pp. 66-93), McGraw-Hill Education, Milano.
- Gandini A., Keeling S., Reviglio U. (2025), “Conceptualising “Algorithmic Public Opinion’: Public Opinion Formation in the Digital Age”, *Dialogues on Digital Society*, 29768640251323147.
- Gran A.B., Booth P., Bucher T. (2021), “To Be or Not to Be Algorithm Aware: a Question of a New Digital Divide?. *Information, Communication & Society*, 24(12), 1779-1796.
- Hargittai E., Gruber J., Djukaric T., Fuchs J., Brombach L. (2020), “Black Box Measures? How to Study People’s Algorithm Skills”, *Information, Communication & Society*, 23(5), 764-775.
- Kamble A., Rao K.B., Navare A. (2026), “Breaking the boredom-fatigue cycle: How algorithmic awareness protects against social media exhaustion”, *Acta Psychologica*, 265, 106768.
- Kitchin R. (2019), “Thinking critically about and researching algorithms, in Beer D. (Ed.), *The Social Power of Algorithms* (pp. 14–29), Routledge, London
- Lupton D., Michael M. (2017), “‘Depends on Who’s Got the Data’: Public Understandings of Personal Digital Dataveillance”, *Surveillance & Society*, 15(2), 254-268.

- Magalhães J.C., Yu J. (2017), “Algorithmic Visibility: Elements for a New Media Visibility Regime”, *European Consortium for Political Research (ECPR)*. Available online: <https://ecpr.eu/Filestore/PaperProposal/e40a7961--0fe3-42a5-8727-f9097552f2fe.pdf> (accessed on 10 May 2019).
- Oeldorf-Hirsch A., Neubaum G. (2023), “Attitudinal and Behavioral Correlates of Algorithmic Awareness among German and US Social Media Users”, *Journal of Computer-Mediated Communication*, 28(5), zmad035.
- Pronzato R., Markham A.N. (2023), “Returning to Critical Pedagogy in a World of Datafication”, *Convergence*, 29(1), 97-115.
- Siles I., Valerio-Alfaro L., Meléndez-Moran A. (2024), “Learning to Like TikTok... and Not: Algorithm Awareness as Process”, *New Media & Society*, 26(10), 5702-5718.
- Sujon Z., Dyer H.T., Bonow Soares F. (2025), Social Media and Society: Platforms, Publics, and Anti-Publics. *Social Media+ Society*, 11(3), 20563051251368234
- Swart J. (2021), “Experiencing Algorithms: How Young People Understand, Feel about, and Engage with Algorithmic News Selection on Social Media”, *Social media+ society*, 7(2), 1-11 20563051211008828.
- Timmermans S., Tavory I. (2012), Theory construction in qualitative research: From grounded theory to abductive analysis. *Sociological Theory*, 30(3), 167-186
- Treré E. (2019), *Hybrid Media Activism: Ecologies, Imaginaries, Algorithms*, Routledge, London.
- Velkova J., Kaun A. (2021), “Algorithmic resistance: Media practices and the politics of repair”, *Information, Communication & Society*, 24(4), 523-540
- Ytre-Arne B., Moe H. (2021), “Folk Theories of Algorithms: Understanding Digital Irritation”, *Media, Culture & Society*, 43(5), 807-824.
- Zarouali B., Boerman S.C., de Vreese C.H. (2021), “Is this Recommended by an Algorithm? The Development and Validation of the Algorithmic Media Content Awareness Scale (AMCA-scale)”, *Telematics and Informatics*, 62, 101607, 1-13.
- Zarouali B., Helberger N., De Vreese C.H. (2021), “Investigating Algorithmic Misconceptions in a Media Context: Source of a New Digital Divide?”, *Media and Communication*, 9(4), 134-144.
- Zhao G. (2025), “Boosting Popularity: Folk Theories and Algorithmic Resistance in Visibility Contests in the Comment Sections” *M Big Data & Society*, 12(2), 1-13.

7. “Palestinian cooking is resistance”.

Algorithmic activism e foodie culture nella resistenza digitale palestinese

di *Claudia Cantale*

1. Introduzione

La resistenza digitale palestinese ha assunto, negli ultimi anni, toni sempre meno “pietisti”, adottando invece un registro giocoso e vivace, coerente con la grammatica e le logiche di piattaforme come Instagram e TikTok. La riconfigurazione dei profili degli influencer, che da “creatori di contenuti” diventano attivisti sembra rispondere al potere del capitalismo di piattaforma (van Dijck *et al.*, 2019; Zuboff, 2019). Nel caso della resistenza digitale palestinese, potrebbe trattarsi di una risposta alla *computational propaganda* israeliana (Bonini e Treré, 2024) attraverso l’adozione di strategie, non sempre consapevoli ma certamente situate, di *algorithmic amplification*, *algorithmic evasion* e *algorithmic hijacking*.

Nel caso specifico, questi attivisti sono anche *food influencer*, per i quali la necessità di preservare e trasmettere la tradizione culinaria palestinese diviene una rivendicazione politica.

Nella prima parte del contributo viene quindi presentata una revisione non esaustiva della letteratura sull’attivismo digitale e sulle pratiche di resistenza algoritmica, che incrociano il lavoro relazionale degli influencer e l’affermarsi dell’*influactivism* nel panorama dei movimenti sociali. Nella seconda parte si rintracciano le pratiche di attivismo digitale nella resistenza palestinese, dando spazio al ruolo del cibo come memoria e come strumento di decolonizzazione del discorso mediale. Nella terza parte viene discusso il commento alla ricerca empirica sui *food influencer* di Gaza.

La ricerca esamina quindi il funzionamento della piattaforma Instagram (*follow the media*) insieme all’analisi di tre account palestinesi (*follow the natives*) (Caliandro e Gandini 2019), intesi come “no-actor as performer” (Senft, 2008; Marwick, 2015) e che, per caratteristiche, possono essere definiti *influ-activist*. Di questi sono stati analizzati 177 reel pubblicati nel pe-

riodo compreso tra dicembre 2023 e novembre 2024. La prima fase dell'analisi include la descrizione dei profili e dei contenuti principali dei reel, insieme ai risultati della *Network Analysis* (d'ora in poi NA), che evidenzia le relazioni con altri attivisti e ONG. La seconda fase di analisi ha invece messo in luce il livello di polarizzazione e il tipo di lavoro relazionale svolto dagli influencer.

2. Attivismo digitale: breve revisione della letteratura

Le tecnologie digitali e la rete Internet hanno contribuito a rinvigorire e plasmare la partecipazione politica di gruppi e attori sociali dentro quell'insieme di pratiche che definiremmo attivismo digitale, come evoluzione dell'attivismo e delle mobilitazioni offline.

Howard Rheingold (2002) evidenzia l'importanza dei processi di coordinamento e cooperazione tra i gruppi di persone attraverso l'uso delle tecnologie *mobile*, come tecnologie abilitanti, connesse alla rete internet, per effetto del *perpetual contact* e *constant touch*. Secondo Castells (2017) nel flusso delle reti le idee, le immagini e i messaggi dei movimenti sociali e degli attivisti hanno possibilità di espandersi e accrescersi nonostante le scarse risorse finanziarie o l'assenza di appoggi politici e istituzionali. Nell'evoluzione dell'attivismo digitale, i movimenti *grassroots* remixano contenuti dei media mainstream (Jenkins, 2014), come il caso della maschera di Guy Fawkes utilizzata per le proteste di Occupy, l'adozione dell'estetica *Handmaid's Tale* nei movimenti femministi, la bandiera Jolly Roger di One Piece nelle rivolte in Nepal.

Alla clandestinità che guida e alimenta le logiche di rete dei gruppi antagonisti (Lovink 2022), si affiancano manifestazioni personalizzate e sempre più professionalizzate di attivismo digitale che non possono prescindere dalle logiche delle piattaforme. L'autonarrazione, necessaria all'abbattimento delle barriere personali che permette l'adesione alle cause, ha influito sull'evoluzione di un tipo di attivismo digitale basato sugli hashtag e sulle strategie di *influencer branded*. L'autonarrazione è basata sulla *self presentation* ed economie emotive per consolidare il capitale sociale e personale; si tratta di impegno politico individuale e non la scrittura di una storia collettiva: alla lotta si partecipa per affermare il proprio punto di vista ed attorno ad esso vengono "aggiunti" quanti più seguaci. Per usare una metafora cara a Lovink (2022), il progressivo impaludamento all'interno delle *oasi* delle piattaforme ha finito per piegare l'attivismo dentro il gioco del capitalismo. L'hashtag *activism* aggrega e polarizza i discorsi, per grappoli di dibattito, discussioni, *thread*, favorendo pratiche di reclutamento, *enga-*

gement. L'hashtag *activism* segue le regole della “memetica” con la scrittura di contenuti che puntano alla viralità prevalentemente visiva. Si tratta quindi di omaggi alle infinite variazioni sul tema, in un flusso perenne di imitazioni dello stesso oggetto.

Dal punto di vista dell'influenza digitale, la letteratura sugli influencer evidenzia come questi ultimi hanno supportato cause politiche e sociali e al contempo in che maniera persone attiviste si siano dotate delle medesime strategie degli influencer anche al fine di rendere remunerativa la propria attività “antagonista” sulle e con le piattaforme (Murru *et al.*, 2024a, 2024b; Baratin e Setiffi, 2024). L'influ-activist è un soggetto che, pur non essendo strutturalmente inserito in partiti o organizzazioni politiche tradizionali, utilizza la propria visibilità e il capitale relazionale accumulato online per promuovere cause sociali o politiche, combinando tattiche dell'influencer marketing, attivismo digitale e partecipazione politica personalizzata. Tale figura rappresenta un'evoluzione del cyber-activist, del digital activist, caratterizzata da una forte individualizzazione dell'azione politica e da un uso strategico delle piattaforme come spazi primari di intervento pubblico (Murru *et al.*, 2024). In sintesi, gli attivisti digitali “partecipano” alle regole d'uso delle piattaforme che sono oasi private dove si orienta il dibattito pubblico, in quanto ambienti socio-materiali. L'iniziale entusiasmo con cui erano accolte le tecnologie digitali è crollato con la fine del mito della *net-neutrality* e dell'immaterialità del capitalismo delle piattaforme (Van Dijk *et al.*, 2017). Secondo Bonini e Treré (2024), gli attivisti digitali sono impegnati quotidianamente a contrastare il potere degli algoritmi, creando azioni per “resistere, sovvertire e riutilizzare gli algoritmi per creare immaginari sociali alternativi e promuovere diversi tipi di futuri dei dati” (p. 38). Il concetto di resistenza algoritmica si riferisce all'insieme di pratiche quotidiane con cui gli utenti, spesso marginalizzati, cercano di esercitare una forma di agency all'interno degli ambienti digitali governati da piattaforme e algoritmi. Questa visione alternativa e contro-egemonica rispetto alla narrazione dominante non vede gli utenti completamente soggiogati dal potere computazionale delle piattaforme. In questa prospettiva la *platform moral economy* (Bonini e Treré, 2024), basata su valori neoliberali, meritocratici, individualisti, si contrappone alla *user moral economy*, con un ritorno ai valori di solidarietà, giustizia, cooperazione, più tipica della rete (Lovink, 2022; Terranova, 2024). Dal punto di vista delle pratiche, la resistenza algoritmica opera sia sul “piano strategico”, con *pratiche di resistenza* ben organizzate e consapevoli, che mediante “tattiche quotidiane” spesso non intenzionali. Gli autori propongono una tassonomia di resistenza *con e contro* gli algoritmi che prevede pratiche finalizzate a massimizzare la visibilità dei contenuti politici (es. messaggi, hashtag, campa-

gne) sfruttando le affordance delle piattaforme (*amplification*); pratiche che aggirano attivamente le logiche di censura, shadow banning o downranking imposte dagli algoritmi (*evasion*) e infine pratiche di appropriazione ostile delle logiche algoritmiche, in cui si “sfrutta” il sistema per finalità antagoniste (*hijacking*). Christal Abidin (2021) sottolinea come l’attenzione dalla visibilità e dalla contestazione diretta si sia spostata verso forme più sottili, intime, e frammentarie di agire politico e culturale e relazionale (silosociability), le chiama “*below-the-radar*”. Anche in questo caso è possibile individuare quattro strategie principali:

- *Masking*: vengono usati codici culturali locali, linguaggi criptati o ambigui, meme ambivalenti, affinché il messaggio circoli senza essere facilmente decodificato. Esempio: l’uso di emoji o slang per criticare governi autoritari in modo non tracciabile.
- *Mirroring*: le forme espressive imitano la struttura estetica o discorsiva dominante, ma ne alterano il significato dall’interno. Si tratta di un modo per “infiltrarsi” nei contenuti algoritmicamente premiati pur mantenendo un’intenzionalità critica.
- *Mimicking*: gli attivisti adottano i codici e le affordance delle piattaforme (es. reels, stories, hashtag, challenge) per veicolare contenuti alternativi, rendendo indistinguibili i contenuti di protesta da quelli culturali dominanti. Ciò consente di “scavalcare” i filtri e i ban automatici.
- *Refracting*: pratica che dà il nome al concetto di *refracted publics*, si riferisce alla creazione di spazi paralleli, effimeri e flessibili in cui le comunità vulnerabili si incontrano, si sostengono e si esprimono, pur sapendo di non poter costruire una sfera pubblica stabile. Esempio: l’uso di gruppi chiusi temporanei su Telegram o Discord, o commenti brevi che si “dissolvono”.

In questo contesto si inserisce la presente riflessione, che mira proprio a rintracciare in che maniera avviene l’addomesticamento dei linguaggi delle piattaforme da parte della resistenza digitale palestinese.

3. Attivismo e resistenza digitale nella causa palestinese

La resistenza palestinese ha utilizzato i media sin dalla Prima Intifada come strategia per preservare la propria memoria collettiva e come strumento di contro-narrazione rispetto all’egemonia culturale occidentale. È tuttavia a partire dalla Seconda Intifada (2000-2005) che si assiste a una svolta significativa verso modelli *grassroots* che caratterizzano l’uso dei media digitali (Jenkins, 2014). In tale contesto, l’accesso a Internet nei territori occupati (Nabulsi, 2014) ha rappresentato una forma alternativa di

mobilità, in risposta alle restrizioni imposte alla libertà di movimento fisico, consentendo ciò che Rousselin (2016) definisce un “nuovo territorio immateriale”, capace di riconnettere le diaspore palestinesi.

Dal fallimento delle aspettative generate dagli Accordi di Oslo (1993–1995) all’onda di fiducia alimentata dalle cosiddette “Primavere Arabe” – che ha valorizzato le piattaforme digitali come strumenti di amplificazione, coordinamento e condivisione (Rheingold, 2002) – i movimenti sociali e gli attivisti digitali palestinesi hanno progressivamente acquisito competenze nell’abitare e nel riappropriarsi degli spazi e dei linguaggi digitali. Un esempio noto, in direzione della decolonizzazione della cultura palestinese, è rappresentato dal movimento digitale globale BDS (Boycott, Divestment and Sanctions).

Tuttavia, le piattaforme digitali, in quanto ambienti socio-materiali, si sono rivelate anche dispositivi di controllo repressivo, strumenti di sorveglianza e di intimidazione *top-down*. A partire dal 2015, il governo israeliano ha intensificato la repressione dell’attivismo digitale palestinese, istituendo un’unità informatica incaricata di monitorare e rimuovere contenuti online considerati di “incitamento alla violenza”. L’ambiguità di tale definizione ha consentito di eludere il *three-part test* sulla libertà di espressione delineato dal Comitato delle Nazioni Unite per i Diritti Umani, contribuendo alla progressiva scomparsa dei contenuti palestinesi dalle principali piattaforme.

Le tecnologie digitali si configurano così come strumenti di restrizione della mobilità non solo nello spazio fisico – attraverso checkpoint e sistemi di riconoscimento facciale – ma anche nello spazio digitale, mediante pratiche di censura, sospensione degli account e persino incarcerazione. In questo modo, l’occupazione israeliana si estende anche alla dimensione digitale, e le piattaforme diventano luoghi di repressione e sorveglianza, sostenute da normative opache che producono il cosiddetto *chilling effect* sulle voci dissidenti e sulle persone cittadine palestinesi. La mediatizzazione del conflitto si è trasferita anche nelle sedi legislative: nel 2017 il Parlamento israeliano, dopo aver proposto la cosiddetta *Facebook Bill*, ha approvato una legge che prevede il carcere per coloro che filmano e documentano le azioni in corso da parte dei soldati israeliani. Le restrizioni alla libertà di espressione imposte alle persone palestinesi risultano illegittime e non conformi agli standard internazionali (Taha, 2020)¹.

¹ In riferimento alle limitazioni delle libertà di espressione delle persone palestinesi è possibile leggere il contributo di S. Taha *The cyber occupation of Palestine*; si segnala anche il rapporto di Amnesty del 2017 in cui si fa riferimento alle misure repressive usate dalla ANP. Per quanto riguarda la progressiva limitazione delle voci palestinesi e pro-Palestina, si

Lo smartphone è arma di propaganda costantemente presente nelle operazioni militari nei territori di Gaza e della Cisgiordania (Kuntsman e Stein, 2020; Kirton, 2016). Le IDF (Israel Defense Force) si sono servite dei social network sin dalla loro diffusione per glorificare l'uso della violenza, aumentare la tensione e disumanizzare le vittime palestinesi. A ciò gli attivisti palestinesi hanno risposto attraverso *livestream* su Instagram e Facebook e video di repertorio su YouTube, tentando di costruire repertori visivi collettivi (Papacharissi e de Fatima Oliveira, 2012), come nel caso dell'Operation Cast Lead (2008) o del docufilm premio Oscar *No Other Land*, realizzato dal collettivo israelo-palestinese Basel Adra, Hamdan Ballal, Yuval Abraham e Rachel Szor.

Tuttavia, i racconti documentari delle violenze subite si scontrano con le forme di *militainment* e *porn war* israeliane, ovvero strategie di comunicazione politica che trasformano la sofferenza e le tragedie altrui in pratiche di consumo piacevole (Stahl, 2009; Garduño García, 2024). Nella colonizzazione dei linguaggi social, ad esempio, il ruolo "seduttivo" attribuito alle donne delle IDF costituisce una testimonianza dei presunti valori democratici e progressisti su cui poggerebbe lo Stato di Israele (Brownfield-Stein 2010), nonché una strategia di evangelizzazione simbolica (Garduño García, 2024).

Dopo il 7 ottobre 2023, la propaganda sionista ha utilizzato i social network in maniera programmata, a partire dal Ministero israeliano per gli Affari della Diaspora, che dispone di un'unità speciale dedicata alla strategia sui social media (Hatzinor, o Unità di Influenza) (Schwartz, 2024), con l'obiettivo di incidere sulla politica estera. TikTok è divenuto un territorio chiave nel conflitto 2023-2025 e ha ridefinito le modalità dell'attivismo, incentivando forme di resistenza giocose, ironiche e creative (Abidin, 2020). Le sue architetture, ereditate da YouTube e ibridate con Vine e Tumblr, sono mutate da Instagram, dove le persone attiviste digitali sono costrette a "giocare" secondo regole imposte da soggetti privati, all'interno di una sfera post-pubblica. Divenuta uno dei pochi strumenti per accedere ai territori di Gaza, la piattaforma è stata successivamente vietata dal governo israeliano. Nei Paesi occidentali, le voci *pro-Pal* vengono silenziate o bannate e, dopo l'accordo di acquisizione della piattaforma da parte di diverse società statunitensi (gennaio 2026), lo *shadowbanning* e il *ban* sono divenuti strumenti operativi di censura, come nel caso emblematico della giornalista Bisan Owda.

Con quella che viene definita la seconda Nakba (Khoury, 2025) e l'avvio

consiglia la lettura della rassegna pubblicata da Di Stefano (2025). Per quanto concerne l'opacità degli algoritmi e le policy delle piattaforme in relazione alla causa palestinese si segnalano invece i contributi N. Abokhodair Y. *et al.* (2024) ed Elmimouni *et al.* (2025).

di una nuova fase di mediatizzazione profonda, l'attivismo digitale palestinese ha assunto toni meno "pietisti", ponendosi in aperta contrapposizione al *militainment* e alla *computational propaganda* israeliana. In questo contesto, vengono adottate strategie di *algorithmic amplification*, *algorithmic evasion* e *algorithmic hijacking*. Possiamo dunque definire la resistenza digitale palestinese come una forma di *algorithmic resistance* (Bonini e Treré, 2024)? È possibile che tale resistenza sconfini in tattiche e strategie di marketing ed *engagement* più tipiche degli influencer, fino a configurarsi come *influ-activism*?

4. La cucina palestinese è *resistenza*

L'urgenza di costruire un racconto della routine del pasto quotidiano da parte dei *food influencer* si oppone alle narrazioni prodotte dal collasso del sistema alimentare nella Striscia di Gaza (dicembre 2023-settembre 2025). L'affamare come arma bellica produce due effetti principali: il primo è relativo alla debilitazione fisico e psicologico della popolazione; il secondo punta alla disumanizzazione puntuale attraverso il racconto degli assalti ai camion dei rifornimenti o delle condizioni di miseria vissute dai palestinesi, dipinti come "animaleschi".

La conversione dei profili degli influencer da "creatori di contenuti dell'industria culturale culinaria" a "partecipanti attivi" è avvenuta in conseguenza all'inasprirsi della crisi e, da un punto di vista empirico, è interessante osservare come la grammatica (Gerlitz e Rieder, 2018) e i vernacoli (Burgess, 2006) di Instagram siano stati adattati dai creator come depositi di memoria. Si tratta di processi di appropriazione e addomesticamento (Silverstone, Hirsh 1993), interpretati anche come reazione al "potere delle piattaforme" (van Dijck *et al.*, 2019), governate da logiche capitaliste (Zuboff, 2019), in cui ci troviamo "impantanati" (Lovink, 2022).

Una delle conseguenze della mediatizzazione della gastronomia è proprio la "ribalta" dello chef, uscito dalla cucina per svelare i segreti al suo pubblico immaginato, in ambienti e con posture più o meno casalinghi (Senft, 2008). La disintermediazione tra cibo cucinato e cibo consumato ha prodotto la professionalizzazione degli *amateur* (pro-am) e il divismo degli e delle chef di professione. La ricerca di un posizionamento personale dentro le economie relazionali alimenta le *foodie cultures* nella direzione performativa del cibo, che non è più solo sostentamento e riconoscimento identitario e culturale². La

² Il cibo costituisce un oggetto classico dell'analisi sociologica in quanto dispositivo di produzione di legami simbolici, solidarietà e forme di autorappresentazione. Lévi-Strauss

condivisione dell'*ars culinaria* e delle proprie vite intime permette processi di identificazione e proiezione tra chef *pro-am* e/o chef celebrity e i loro fan. Va sottolineato che il mondo professionalizzato dei food influencer e del divismo culinario è dominato da persone etero, abili e bianche, che contribuiscono a consolidare l'egemonia culturale occidentale e più raramente a problematizzarla (Salvio, 2011; Dejmanee, 2023). La *foodie culture* è, in sostanza, colonizzata dalla cultura euro-americana, spesso basata su appropriazioni delle cucine "altre", trasformate in immaginari "etnici", "sani" e "tradizionali", attraverso cui è possibile esercitare il controllo politico del corpo.

In questo senso, la *foodie culture* può essere intesa come una forma di orientalismo digitale: l'astrattismo dei regimi del capitalismo digitale tende a silenziare e appiattire le pratiche alimentari delle minoranze. Secondo Beck (2012), questo processo dissolve i legami di socializzazione familiare e collettivi, rafforzando l'individualizzazione delle prassi religiose, non necessariamente legate a luoghi, memorie e istituzioni sacre. Dal saccheggio e dalla mercificazione delle culture alimentari deriva la necessità di preservare le tradizioni anche attraverso la funzione mnemonica dei media.

Nel caso della rivendicazione del riconoscimento delle origini di molte ricette tradizionali palestinesi si tratta, appunto, di una questione politica (Kassis, 2017). Come ricorda Nasser (2024a, 2024b), i cookbook best-seller israeliani, eliminando i riferimenti alle origini palestinesi delle ricette, si configurano come dispositivi di normalizzazione della colonizzazione e della guerra, presentando l'occupazione e la violenza strutturale come sfondo neutro o come semplice questione di convivenza culturale. Allo stesso tempo, essi legittimano una riconfigurazione capitalistica del territorio, in cui lo sfruttamento intensivo dei suoli e la distruzione dei sistemi agricoli indigeni sono narrati come progresso, abbondanza e modernizzazione.

Il cibo, vettore di resistenza e custode di miti (Nabulsi, 2014; Kasturi, 2023), assume così la funzione di strumento apparentemente de-politicizzato, capace di bypassare le prassi censorie degli algoritmi, fondamentali anche nell'escalation del conflitto israelo-palestinese (2023-2026).

(1964) legge nelle pratiche culinarie una strutturazione della socialità, inscritta nelle procedure implicite dell'atto del cucinare; Simmel (1910) ne evidenzia il valore relazionale e la connessione con le dinamiche di classe. In quanto bene simbolico, il cibo diviene per Bourdieu (1979) un marcatore di distinzione, espressione dell'*habitus* più che di una scelta individuale. Nei banchetti e nei rituali collettivi, infine, esso assume una dimensione sacrale (Durkheim), sospendendo temporaneamente l'ordine profano e riaffermando le norme della vita comunitaria. Per una rassegna sulla relazione tra cibo e social media si segnalano le letture di Guido Nicolosi (2007); Luisa Stagi (2016); Tito Vagni (2017); Antonia Cava (2018).

5. Il disegno della ricerca

Il contesto di precarizzazione provocato dall'inasprirsi del conflitto ha visto la conversione dei profili degli influencer da digital creator di contenuti per l'industria culturale a partecipanti attivi di movimenti politici, come influ-activist. Dal punto di vista empirico quindi si tratta di osservare come vengono adattate la grammatica (Gerlitz and Rieder, 2018) e i vernacoli (Burgess, 2006) di Instagram alla nuova veste dei food-influencer.

La ricerca, quindi, prende in esame il funzionamento della piattaforma Instagram (follow the media)³ insieme all'analisi di tre account palestinesi (follow the native), @hamadashoo (569.000 follower), @basma_shahla (532.000 follower) e @dr.mo5tar_ (62.000 follower).

I profili sono stati selezionati seguendo i seguenti criteri:

1. *Micro-celebrità*: per numero di follower rientrano nella classificazione dei micro o mid - influencer (50k-500k). Theresa Senft (2013) parla di microcelebrità nel senso di *no-actor as performer*, ovvero di una persona che viene seguita non per quello che mostra di essere ma per quello che fa o sa fare (Marwick, 2015)⁴.
2. Tipo di profilo: gli account selezionati sono verificati e hanno un background da digital creator come micro (nano)-influencer, conoscono quindi la logica delle piattaforme e la *influencer culture* come insider. Si tratta di influencer verticali, ovvero che hanno il cibo come tema principale.
3. *Level of activities*: gli account selezionati hanno un ritmo di pubblicazione costante, con uno o due contenuti settimanali. I reel sono muti, sottotitolati in inglese, e presentano la caption in doppia lingua, inglese e araba.

Con un *tool scraper* (Phantombuster) adatto a Instagram sono state estratte informazioni relative a 205 *reel* postati dai tre creators nel periodo compreso tra dicembre 2023 e novembre 2024, ridotti poi a 177. L'analisi include la descrizione dei profili e l'analisi visuale dei *reel*. Le categorie

³ È doveroso sottolineare che la costruzione del caso di studio è iniziata a partire dal mio account personale di Instagram che dall'ottobre del 2023 segue gli hashtag #gaza #freepalestine.

⁴ Qui facciamo riferimento alle Linee Guida individuate dai tavoli tecnici di AGCOM per le categorie definitorie di digital creators e influencer: si vedano i testi «Linee-guida volte a garantire il rispetto delle disposizioni del Testo unico da parte degli influencer e istituzione di un apposito Tavolo tecnico», e «Tavolo tecnico per definire le ulteriori misure per il rispetto delle previsioni del testo unico da parte degli influencer» pubblicato a giugno del 2026 <https://www.agcom.it/tavolo-tecnico-definire-le-ulteriori-misure-il-rispetto-delle-previsioni-del-testo-unico-da-parte>.

analitiche riportate nella tavola 1 sono state elaborate per definire in che maniera i tre creators hanno adattato i linguaggi di Instagram per poterne sfruttare appieno le possibilità offerte. Per quanto riguarda il contenuto, per esempio sono stati distinti i tipi di contenuti, personali, di intrattenimento e documentaristici. Nella costruzione della matrice analitica si è tenuto conto della distinzione dei contenuti personali in *lifestyle* (Bennet e Segerberg, 2013; Baratin e Seffiti, 2024) e *intimate* per chiarire il posizionamento dei creators come attori influ-attivisti (Mazzoni e Ciaglia, 2015; Murru *et al.*, 2024). La doppia concettualizzazione proposta da Vagni (2017) di *food porn* e *food media* evidenzia gli aspetti di pornificazione e dello sguardo focalizzato sul cibo (Barthes, 1980). Si tratta di usare quindi l'analisi semiotica dello stile drammaturgico, che prende in considerazione qual è l'oggetto/soggetto del desiderio dell'occhio-camera.

Ne deriva la classificazione dello stile, basata sulla tipologia di video e di ripresa sul modello di Cervi, Tejedor e Blesa (2023) sulla varietà e le significazioni delle riprese presenti su TikTok (Abidin, 2020; Rao *et al.*, 2020). La presenza di attori riconoscibili (come bambini, giornalisti, celebrità), insieme allo scenario in cui la narrazione si iscrive (campi profughi, veicoli in movimento, paesi occidentali), servono a rintracciare eventuali pratiche cosmetiche (Boccia Artieri *et al.*, 2018), funzionali alla NA. La classificazione dei generi dentro cui si inseriscono i format più ricorrenti – che tiene conto del ruolo di elementi come la musica e i filtri, quali potenziali aggregatori di viralità (Vizcaino-Verdú e Abidin, 2022), e di elementi creativi come l'ASMR (Autonomous Sensory Meridian Response)⁵, le emoji, l'algospeak, gli hashtag activism – serve a descrivere come i creators adottano il linguaggio di piattaforma (Caliandro *et al.*, 2024; Burgess, 2006) per attuare forme di resistenze algoritmiche (Bonini e Trerè, 2024).

⁵ Con L'ASMR, Autonomous Sensory Meridian Response, si intende la risposta sensoriale provocata da uno stimolo che si percepisce come particolarmente piacevole, accompagnato da un senso di rilassamento, formicolio e benessere generale. Tecnicamente si tratta di una forma di piacere che un individuo riesce a procurarsi autonomamente. Il termine ASMR è divenuto mainstream con la sua diffusione sui social, trasformandosi in un genere per la produzione di contenuti particolarmente attrattivi.

Tav. 1 - Griglia di analisi dei contenuti visuali

Dimensioni	Indicatori
<i>Genere / Tipologia di contenuto</i>	Video Duetto Challenge Carosello Meme GIF
<i>Formato</i>	Solo video Video con effetti Video con effetti e musica Video con musica Video con testo Video con testo e musica Video con effetti, musica e testo
<i>Elementi creativi</i>	ASMR (risposta sensoriale autonoma) Emoji Attivismo tramite hashtag Algospeak (linguaggio adattato agli algoritmi) Didascalia
<i>Tono edonico / emotivo</i>	Gioia (allegro, felice, soddisfatto) Tristezza (depresso, triste, abbattuto) Rabbia (irritato, infastidito, furioso, arrabbiato) Paura (ansioso, teso, inquieto, nervoso) Disgusto (sgradevolezza, rifiuto)
<i>Fonte</i>	Partiti politici Istituzioni Media Utenti ONG
<i>Contenuto</i>	Personale Intrattenimento alimentare (food entertainment) Documentaristico
<i>Stile – tipo di inquadratura</i>	Primo piano (Close Up – CU) Campo lungo / figura intera (Full Shot – FS) Mezzo primo piano (Medium Close Up – MCU) Campo medio (Medium Shot – MS) Inquadratura a due soggetti (Two Shot) Campo largo (Wide Shot) Punto di vista soggettivo (POV) Zoom
<i>Stile – tipo di ripresa</i>	Selfie Amatoriale / fatto in casa (Homemade) Professionale Generato con intelligenza artificiale

<i>Attori principali</i>	Attivisti Celebrità Influencer (creator di contenuti) Politici Rifugiati Bambini Volontari
<i>Paesaggi principali</i>	Campo profughi Contesto urbano Edifici / architetture Ambiente naturale Veicoli Paesaggio occidentale
<i>Interazione</i>	Partecipazione e mobilitazione Raccolta fondi Informazione e promozione

Per concludere l'analisi sui *testi* proposti dai creators si è proceduto all'analisi del contenuto estraendo i temi principali mediante una matrice *TF-IDF*⁶ delle *captions*. La frequenza delle parole non ci dice tutto di un testo (Ciotti, 2017); tuttavia la rappresentazione grafica e computazionale della wordcloud consente di trovare chiavi interpretative non sempre accessibili mediante la *close reading* (Moretti, 2013; Esposito, 2023).

La NA, realizzata a partire dai dati estratti da *tags* (come collaborazioni al post) e dalle *mentions* presenti nelle *captions* mira a descrivere le collaborazioni esistenti tra influencer, attivisti, ONG o legacy media, evidenziando il contributo dell'influenza reciproca (Barabasi, 2002; Rogers, 2013) *pro-Pal*⁷.

Per l'analisi dei commenti si è adottato un metodo fondato su ChatGPT e *prompt engineering*. L'architettura del progetto prevede l'adozione di linee guida, di letteratura teorica metodologica (Geertz 1973, 2008; Montesperelli, 2016), e sulla colonizzazione e decolonizzazione della cucina palestinese. Per la realizzazione del *code-book* (Tav. 2) è stato fornito quindi un elenco di istruzioni, il contesto di riferimento, alcune fonti ed esempi di commenti come guida alla definizione delle categorie, e infine sono state effettuate alcune sistemazioni.

⁶ La matrice TF-IDF (Term Frequency-Inverse Document Frequency) serve a misurare l'importanza di alcuni termini nel contesto di una collezione testuale. La loro visualizzazione viene spesso realizzata attraverso le wordcloud.

⁷ Da questo momento intendiamo con *pro-Pal* coloro che esplicitano il proprio sostegno al popolo palestinese, senza entrare nel merito del dibattito sul movimento.

<i>emotive ed empatia</i>	forti reazioni emotive, con commenti che esprimono empatia e coinvolgimento emotivo. Viene evidenziato il senso di impotenza di fronte alla sofferenza dei bambini, ma anche la speranza riposta in loro.	1=condanna 2 = colpevolizzazione 3 = solidarietà [Sono considerate solo le espressioni che fanno esplicito riferimento a emozioni (pianto, gioia, paura, speranza)].	<ul style="list-style-type: none"> • “This brought tears to my eyes 😭” • “Missing a dead person is the saddest thing... life is cruel and short” • “I hope each and every one of these children gets to travel everywhere their heart desires”
<i>Entusiasmo culinario</i>	I commenti fanno riferimento al cibo e alla tradizione palestinese. Gli utenti esprimono giudizi sulle ricette e, più in generale, sul progetto gastronomico.	0=neutro 1=condanna 2 = colpevolizzazione 3 = solidarietà [Sono considerati solo i commenti che menzionano esplicitamente la ricetta o il cibo].	<ul style="list-style-type: none"> • “People in a genocide be cooking better than Indians 🍔🍷” • “It looks delicious! 🍰🍷🍩” • “Oh dear God that looks so good. Tiramisu is my all time favorite dessert food and I can tell you enough how much I want to eat that right now, it looks so delicious 😊” • “Grateful for this reel ❤️ hearing the laughter is wonderful and tasting this is delicious this morning may your cooking be a sparkling of goodness always. Love and Freedom ❤️”
<i>Reazioni visive basate su emoji</i>	I commenti sono composti esclusivamente da emoji, senza testo verbale.	0=assente 1 = presente [Se presenti solo emoji → 1]	<ul style="list-style-type: none"> • “❤️❤️❤️❤️❤️❤️❤️❤️❤️❤️❤️”

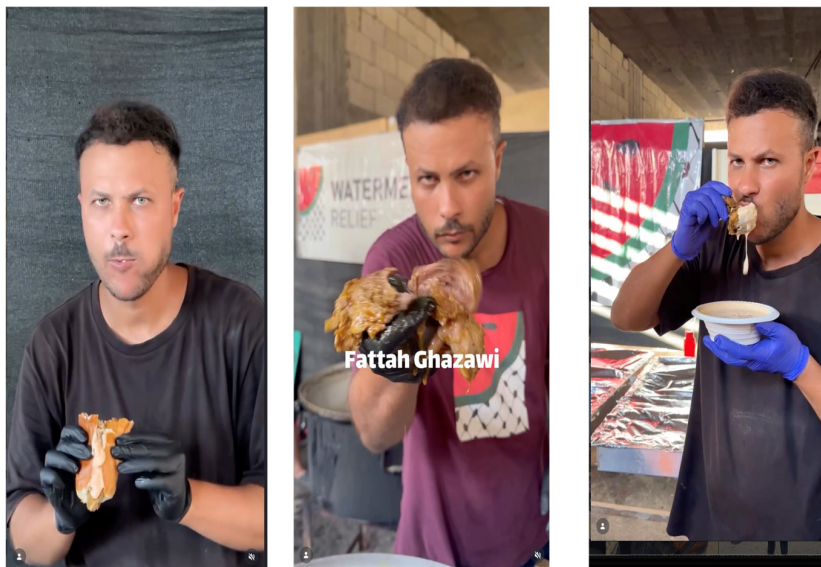
Altro	<p>Commenti non pertinenti al contenuto principale, solitamente legati alla cultura pop (canzoni, film, serie TV, riferimenti algoritmici).</p>	<p>0=assente; 1 = presente</p> <p>[Se presenti riferimenti estranei al contenuto → 1]</p>	<ul style="list-style-type: none"> • “<i>May they and their family find peace and safety through this injustice pop culture references for the algorithm (more u can think of to help!): bridgerton, taylor swift, eloise, avatar. the last airbender, atla, zuko, azula, kendrick lamar, colin may you and your family find peace and safety through this injustice</i>”
-------	---	---	--

6. Riflessioni conclusive

Dall’analisi visuale sono emerse alcune differenze salienti nelle scelte vernacolari ed estetiche dei creator, riconducibili a format di genere

I reel di *Hamada* e *Dr. Mo5tar_* sono video di intrattenimento il cui format prevede la realizzazione di una ricetta (food entertainment), in cui vengono privilegiate riprese *close-up* e *medium close-up*, focalizzate sui corpi e sui volti dei cuochi. Gli ingredienti vengono zoomati durante l’esecuzione e nella composizione del piatto finale. Prima che il cibo venga distribuito ai bambini e alle bambine (*activist/kids – content creator/kids*), *Hamada* e *Dr. Mo5tar_* spezzano, assaggiano, massaggiando, sfregano i coltelli sulle panature per verificarne il “fattore crunch”, ottenendo un effetto di erotizzazione della scena. Il piacere deriva dalla procedura e dalle modalità di esecuzione della ricetta, come forma di ri-mediazione della sitofilia (Vanni, 2017). In piena coerenza con l’estetica #foodporn, *Hamada_shoo* massaggia e “sculaccia” panificati e carne, guarda in camera mentre assaggia i suoi piatti, dimostrando una padronanza ancestrale del fuoco (Figura 1).

Fig. 1 - Selezione di screenshot dall'account di Instagram di Hamada_shoo



Nel caso dei *tent chef creator*, la componente scenografica prevalente è la cucina da campo (*refugee camp*), da cui scompaiono i riferimenti a contesti urbani o abitativi: solo raramente vengono ripresi edifici e non vi sono tracce evidenti del conflitto in corso. Si direbbe che si tratti di un'operazione cosmetica di sottrazione (Boccia Artieri *et al.*, 2018) a bassa intensità, in cui i contenuti poco graditi agli algoritmi vengono esclusi dal discorso. Vi è da dire che questa sottrazione non va interpretata necessariamente come una strategia pienamente consapevole, ma come una pratica situata che emerge dall'interazione tra *agency* dei creator e vincoli strutturali imposti dalla piattaforma. Sempre in questa direzione è opportuno sottolineare come l'impiego del ronzio dei droni dell'esercito israeliano all'interno del format ASMR, o di video con effetti musicali e sonori associati al borbottio e allo sfrigolio delle cucine, sia da considerarsi un'applicazione 'spontanea' di tattiche di *masking* e *mirroring* del linguaggio creativo nella più ampia strategia di resistenza digitale *below the radar* adottata dai tre creator.

Dr. Mo5tar_ predilige il duetto e la challenge (*mimicking*), ottenendo come esito la rivisitazione di piatti provenienti da cucine estere, adattati agli ingredienti disponibili nella cucina da campo e pensati per un numero significativo di persone (Figura 2). L'adozione del format della challenge,

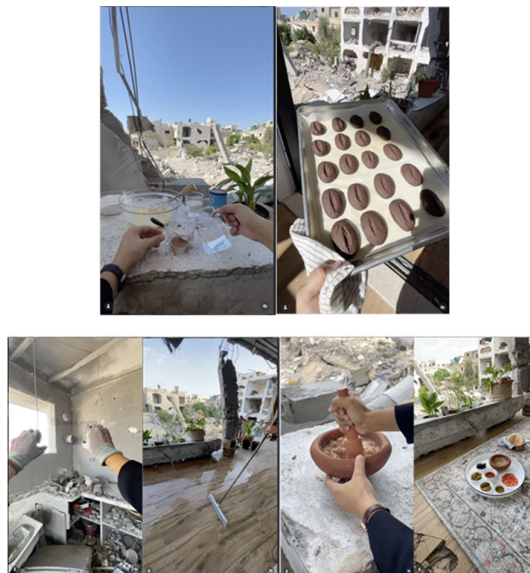
nel caso di *Hamada* e *Dr. Mo5tar_* conferisce un tono allegro che si contrappone alle tinte tragiche dei racconti di guerra in generale, e in particolare alla narrazione del conflitto a Gaza. In questo modo il contenuto appare depoliticizzato e quindi poco problematico, risultando maggiormente coerente con le logiche di amplificazione e risonanza gradite agli algoritmi e producendo un effetto positivo in termini di connessioni affettive (Papacharissi, 2015).

Fig. 2 – Selezione di screenshot dall'account di Instagram con duetti e challenge di *Dr. Mo5tar_*



Il regime estetico di *Basma* (Figura 3) è costruito attraverso la prospettiva in soggettiva della macchina da presa (POV): la camera è rivolta sulle sue mani che riassettano, sistemano e puliscono (lifestyle) la location destinata ad accogliere il piatto finito, consumato dalla stessa protagonista. Lo schema proposto è quello del *day in the life*. Assumendo il punto di vista della creator, i racconti assumono un tono personale e intimo (*personal/intimate*). L'effetto ottenuto è simile a quello del *cleaning house aesthetic*, in cui pulizia e ordine diventano strumenti di resilienza. Casey e Littler (2021) descrivono il *cleanfluencing* come una forma di terapia contro l'ansia, in cui pulire e organizzare offre anche piacere estetico. Le *cleaninfluencer* sono percepite come “persone pulite” in virtù di uno stile di vita che riflette i *topoi* romantici proposti da media genderizzati al femminile. Il racconto visivo di *Basma* è centrato sulla cura, e in particolare sulla cura dei dettagli, sublimata nella *mise en place*, in perfetta sintonia con la narrazione della cucina al femminile (Stagi, 2016; Cava, 2018).

Fig. 3 - Selezione di screenshot dall'account di Instagram con sequenza di mise en place di Basma



Basma imita i codici linguistici delle piattaforme assumendo la postura delle donne il cui ruolo è quello della cura domestica, rimettendo al centro la casa e l'abitare come spazi politicizzati per rivendicare il diritto alle “chiavi di casa”, tema storicamente legato alla questione palestinese sin dalla Nakba del 1948. L'adozione di questi codici non va letta come una naturalizzazione dei ruoli di genere, bensì come un riuso strategico di immaginari femminilizzati già sedimentati nei media di piattaforma. Alla segregazione domestica (Negra, 2009) e al *displacement* forzato, contrappone la routine culinaria come spazio sicuro in cui continuare ad avere cura di sé, inteso come approdo rituale. Basma incarna l'ideale della donna che, con raffinata delicatezza, mostra le proprie emozioni alla telecamera: è glamour senza ostentazione, indossa la keffiah, abiti tradizionali e accessori di lusso occidentali, beve tè matcha ed è impegnata in una personale ricerca di equilibrio fondata su sguardo interiore, gentilezza, gratitudine e gioiosa consapevolezza (Casey, 2025).

In questo caso la *word cloud* (Grafico 1) è utilizzata come strumento esplorativo, utile a individuare ricorrenze tematiche, e non come tecnica inferenziale autonoma, ci aiuta a orientarci nell'interpretazione della relazione tra posizionamento da *food influencer* e quello da *influ-activist*.

“brandizzate” con il logo dell’organizzazione grassroots Watermelon Relief. Considerando la dimensione dei nodi, Hamada si colloca allo stesso livello di giornalisti come Wissam Nassar, photographer (160k), e Ahmed Eldin, journalist (1.1M). Inoltre, Hamada cura attivamente la propria comunità (*the hamada sho family*) anche attraverso canali alternativi come Telegram e offre servizi a pagamento, tra cui lezioni di lingua araba.

Fig. 5 – Network Analysis – tags Instagram. Grafo elaborato con Gephy



Un ulteriore elemento emerso dalla mappatura della rete è la geminazione di account di backup (ad esempio @EveryDayPalestine e @everydaypalestine2; @jenan.matari e @jenanmatari), come nel caso dello stesso Dr. Mo5tar , interpretabile come reazione alle politiche di *ban* e come tattica di “rifrazione”.

Infine, anche la risposta dell’audience di Instagram fornisce indicazioni sulle tattiche cosmetiche dei creator e delle loro comunità, seppur transitorie. I commenti di segno positivo sembrano essere una conseguenza della moderazione censoria, che limita commenti e risposte indesiderate tali da generare thread di conversazioni polarizzate di segno negativo. Nella maggioranza dei casi (tabella 1) si tratta di commenti di sostegno e solidarietà, carichi di pathos ed empatia, in cui la componente arabofobica o islamofobica risulta pressoché assente. Gli utenti ricorrono frequentemente a linguaggi mimetici della piattaforma, come l’uso tattico di emoji, *algospeak*, strofe di canzoni di Billie Eilish o paragrafi di Wikipedia, che svolgono una funzione di amplificazione migliorando la performance in termini di engagement e permettendo ad hashtag come #freepalestine di rendersi invisibili ai *ban* algoritmici di Instagram.

Tab. 1 – Dimensioni di analisi, distribuzione in %

Tipo di comment	Valore assoluto	Percentuale
Solidarietà e supporto per la Palestina	346	38,4%
Benedizioni e preghiere religiose	186	20,7%
Gratitudine e apprezzamento	128	14,2%
Risposte emotive ed empatia	21	2,3%
Entusiasmo culinario	14	1,6%
Reazioni visive e basate su emoji	196	21,8%
Altro	9	1,0%
<i>Totale</i>	<i>900</i>	<i>100,0%</i>

In conclusione, i profili degli influencer analizzati mantengono la natura apparentemente frivola e narcisistica dell’auto-narrazione, necessaria per catturare l’attenzione, al fine di costruire un discorso orientato, fortemente empatico e affettivo. È possibile definirli una ri-mediazione “giocosa” dei diari o dei blog di guerra. Condividendo e archiviando il proprio ricordo biografico, l’*influ-activist* assume il ruolo di delegato del sapere che, nel caso palestinese, contribuisce a conservare e modellare la memoria collettiva (Assmann, 1997). Instagram diventa così un archivio delle pratiche e della resistenza in uno scenario mediatico dominato dall’assenza o dall’annientamento delle fonti giornalistiche, in presenza di perdite materiali e immateriali ingenti. Il controllo e la gestione dei commenti, così come l’adesione ai *vernacular platform* (uso di musica, ASMR, parafilia, duetti, challenge), rientrano nelle strategie di resistenza algoritmica (Qr1) attraverso tattiche di *algorithmic evasion* e *algorithmic amplification*. In questa prospettiva, il foodismo e la *foodie culture* (Stagi, 2016; Cava, 2018) vengono ri-significati come progetto di attivismo digitale. In questo senso, la decolonizzazione delle piattaforme non va intesa come trasformazione strutturale dell’infrastruttura, ma come insieme di micro-pratiche vernacolari di riappropriazione simbolica.

La ricerca si configura come una soglia di ingresso ai peculiari percorsi di resistenza digitale e algoritmica palestinese, evidenziandone al contempo alcuni limiti strutturali, in primo luogo perché il disegno di ricerca non consente di misurarne gli effetti in termini di posizionamento da parte delle community rispetto a quanto accade in Oriente. Vale a dire che il ruolo svolto dalle strategie di resistenza digitale nel creare effetti di cumulazione e consonanza potenzialmente capaci di determinare una momentanea interruzione della spirale del silenzio e di concorrere alle mobilitazioni di piazza del 2025 rimane sullo sfondo dell’analisi.

Un approccio più ecosistemico risulterebbe funzionale a cogliere, ad esempio, in che modo i network di *influ-activist* contribuiscono a processi di *pollination*, nella misura in cui “spargono” la voce, ma anche correndo il

rischio di alimentare, paradossalmente, forme di rumore e di *pollution* informativa.

In secondo luogo, un ulteriore limite è riconducibile allo scarto temporale tra la raccolta e l'analisi dei dati e il momento della pubblicazione, periodo segnato da una rapida evoluzione degli eventi con conseguenze rilevanti in termini tecnologiche e politiche.

La trasformazione accelerata del contesto solleva interrogativi sulla natura epistemica della riflessività nella pratica della ricerca, in particolare laddove esposto alle pressioni delle *rabbit holes* algoritmiche, può condurre a un uso degli stessi strumenti di indagine anche come dispositivi di adattamento cognitivo.

Riferimenti bibliografici

- Abidin C. (2020), "From "Networked Publics" to "Refracted Publics": A Companion Framework for Researching "Below the Radar" Studies", *Social Media + Society*, 7(1), <https://doi.org/10.1177/2056305120984458>
- Abidin C. (2021), "Mapping Internet Celebrity on TikTok: Exploring Attention Economies and Visibility Labours", *Cultural Science Journal*, 12(1), 77–104.
- Abokhodair N., Skop Y., Rüller S., Aal K., Elmimouni H. (2024), "Opaque Algorithms, Transparent Biases: Automated Content Moderation during the Sheikh Jarrah Crisis", *First Monday*, 29(4) <https://doi.org/10.5210/fm.v29i4.13620>
- AGCOM (2024a), *Linee-guida volte a garantire il rispetto delle disposizioni del testo unico da parte degli influencer e istituzione di un apposito tavolo tecnico* <https://www.agcom.it/provvedimenti/delibera-7-24-cons> (16 gennaio 2024).
- AGCOM (2024b), *Tavolo tecnico per definire le ulteriori misure per il rispetto delle previsioni del testo unico da parte degli influencer* <https://www.agcom.it/tavolo-tecnico-definire-le-ulteriori-misure-il-rispetto-delle-previsioni-del-testo-unico-da-parte> (16 gennaio 2024).
- Amnesty International (2017), *State of Palestine: Alarming attack on freedom of expression* (Index No. MDE 15/6983/2017).
- Assmann J. (1997), *La memoria culturale: Scrittura, ricordo e identità politica nelle grandi civiltà antiche*, Einaudi, Torino.
- Barabási A.L. (2002), *Link: La scienza delle reti*, Einaudi, Torino.
- Baratin A., Setiffi F. (2024), "Lifestyle Politics Narratives without Full Political self-Identification: A Qualitative Study of Italian Eco-influ-activists Active on Instagram", *Mediascapes Journal*, 24(2), 114–133.
- Beck U. (2012), *Il Dio personale: La nascita della religiosità secolare*, Laterza, Bari.
- Bennett W.L., Segerberg A. (2013), *The Logic of Connective Action: Digital Media and the Personalization of Contentious Politics*, Cambridge University Press, Cambridge.
- Boccia Artieri G., Gemini L., Pasquali F., Carlo S., Farci M., Pedroni M. (2018),

- Fenomenologia dei social network: Presenza, relazioni e consumi mediali degli italiani online*, Guerini e Associati, Milano.
- Bonini T., Treré E. (2024), *Algorithms of Resistance: The Everyday Fight against Platform Power*, MIT Press, Cambridge.
- Brownfield-Stein C. (2010), “Visual Representations of IDF Women Soldiers and ‘Civil-Militarism’ in Israel”, in Sheffer G., Barak O. (Eds.) *Militarism and Israeli Society*, (pp. 304-328), Indiana University Press, Bloomington,
- Burgess J. (2006), “Hearing Ordinary Voices: Cultural Studies, Vernacular Creativity and Digital Storytelling”, *Continuum*, 20(2), 201–214.
- Burkhardt S., Rieder B. (2024), “Foundation Models are Platform Models: Prompting and the Political Economy of AI”, *Big Data & Society*, 20(2) <https://doi.org/10.1177/20539517241247839>
- Caliandro A., Gandini A., Bainotti L., Anselmi G. (2024), “The Platformization of Consumer Culture: A Theoretical Framework”, *Marketing Theory*, 24(1), 3–21.
- Calvo P., García C.S. (2025), “Generative AI and Democracy: Risks and Impacts behind the Production of Synthetic Public Opinion”, *SSRN*. 01/2025. SSRN 4911710
- Cardon D. (2016), *Che cosa sognano gli algoritmi*, Mondadori, Milano.
- Casey E. (2025), *The Return of the Housewife: Why Women are still Cleaning up*, Manchester University Press, Manchester.
- Casey E., Littler J. (2022), “Mrs Hinch, the Rise of the Cleanfluencer and the Neoliberal Refashioning of Housework”, *The Sociological Review*, 70(3), 489–505.
- Castells M. (2017), *Comunicazione e potere*, Università Bocconi Editore, Milano.
- Cava A. (2018), *#Foodpeople: Itinerari mediali e paesaggi gastronomici contemporanei*, Aracne, Roma.
- Cervi L., Tejedor S., García Blesa F. (2023), “TikTok and Political Communication: The latest Frontier of Politainment?”, *Media and Communication*, 11(2), 203–217.
- Ciotti F. (2017), “What’s in a Topic Model?”, *Testo & Senso*, (18),
- De Blasio E., Selva D., Sorice M. (2022), “Il dibattito sul DDL Zan e la post-sfera pubblica italiana”, *Mediascapes Journal*, 19(1), 89–112.
- Dejmanee T. (2019), “The Food Network’s Heartland Kitchens: Cooking up Neo-conservative Comfort in the United States”, *Critical Studies in Television*, 14(1), 74–89.
- Di Stefano S. (2025), “Silencing Palestinian Voices: On Freedom of Expression and Gaza”, *Netherlands Quarterly of Human Rights*, 43(1), 3–11.
- Durkheim É., Rosati M. (Eds.) (2013), *Le forme elementari della vita religiosa: Il sistema totemico in Australia*, Mimesis, Milano.
- Elmimouni H., Rüller S., Aal K., Skop Y., Abokhodair N., Wulf V., Tolmie P. (2025), “Exploring algorithmic resistance: Responses to social media censorship in activism”, *Proceedings of the ACM on Human-Computer Interaction*, 9(2), 1–24.
- Espósito E. (2022), *Comunicazione artificiale: Come gli algoritmi producono intelligenza sociale*, EGEA, Roma.
- Floridi L. (2022), *Etica dell’intelligenza artificiale*, Raffaello Cortina, Milano.

- Garduño García M. (2024), “La política de seducción”, *Estudios de Asia y África*, 59(3), 25–49.
- Geertz C. (1973), *The interpretation of cultures*, Basic Books, New York.
- Geertz, C. (2008), “Thick description”, in *The cultural geography reader* (pp. 41–51), Routledge, London.
- Gerlitz C., Rieder B. (2018), “Tweets are not created equal”, *International Journal of Communication*, 12, 528–547.
- Highfield T., Leaver T. (2015), “A methodology for mapping Instagram hashtags”, *First Monday*, 20(1).
- Jenkins H. (2014), *Cultura convergente*, Maggioli, Santarcangelo di Romagna.
- Kassis R. (2017), *The Palestinian table*, Phaidon, New York.
- Khoury N. (2025), “The meanings of a second Nakba”, *Journal of Genocide Research*, 28 (2), 460–475.
- Kirton A. (2016), “Online engagements: War and social media”, in *The Palgrave handbook of criminology and war* (pp. 407–424), Palgrave Macmillan, London.
- Kuntsman A., Stein R.L. (2020), *Digital militarism: Israel’s occupation in the social media age*, Stanford University Press, Stanford.
- Lovink G. (2022), *Le paludi della piattaforma*, Nero, Roma.
- Manovich L. (2020), *L’estetica dell’intelligenza artificiale*, Luca Sossella, Roma.
- Marwick A.E. (2015), Instafame. *Public Culture*, 27(1), 137–160.
- Montesperelli P. (2016), “Comunicazione, metodo e interpretazione”, *Sociologia della comunicazione*, 50/2015, 19-1.
- Moretti F. (2013), *Distant reading*, Verso, London-New York.
- Murru M.F., Pedroni M., Tosoni S. (2024a), “Influ-activism”, *Mediascapes Journal*, (24), 1–20.
- Murru, M. F., Pedroni, M., Tosoni, S. (2024b), “Tra influence culture e attivismo digitale”, *Quaderni di Teoria Sociale*, 2, 13–44.
- Nabulsi M. (2014), “Hungry for freedom”, in Herrera L., Sakr R. (Eds.), *Wired citizenship* (pp. 105–120), Routledge, London.
- Nasser C.D. (2024a), “Food and the transmission of culture”, *Jerusalem Quarterly*, 99, 7. <https://doi.org/10.70190/jq.199.p6>
- Nasser C.D. (2024b), “Palestinian food: Commensality and cultural resistance”, in *Routledge handbook on Palestine* (pp. 507–522), Routledge, London.
- Nicolosi G. (2007), *Lost food: Comunicazione e cibo nella società ortoressica*. Edit Press, Firenze.
- Papacharissi Z. (2015), *Affective Publics*, Oxford University Press, Oxford.
- Papacharissi Z., Fatima Oliveira M. (2012), Affective News and Networked Publics, *Journal of Communication*, 62(2), 266–282.
- Rheingold H. (2002), *Smart mobs: the next social revolution*. Routledge, London.
- Rousselin M. (2016), “Modern Communication Technologies and the Extension of the Territory of Struggle”, *New Media & Society*, 18(7), 1201–1218.
- Salvio P.M. (2012), “Dishing it out”, *Gastronomica*, 12(3), 31–39.
- Senft T.M. (2013), “Microcelebrity and the Branded self”, in Hartley J., Burgess J., Bruns A. (Eds.), *A Companion to New Media Dynamics* (pp. 346–354), Blackwell Publishing, Hoboken.

- Silverstone R., Hirsch E. (1993), “Information and Communication Technologies and the Moral Economy of the Household”, in Silverstone R., Hirsch E. (Eds.), *Consuming Technologies* (pp. 25-40), Routledge, London
- Srnicek N. (2017), *Platform Capitalism*, Polity Press, Cambridge.
- Stagi L. (2016), *Food Porn*. EGEA, Roma.
- Stahl R. (2009), *Militainment, Inc.* Routledge, London.
- Sweeney-Romero K.M. (2022), “Wellness TikTok”, in Boffone T. (Eds.), *TikTok Cultures in the United States* (pp. 108-116), Taylor & Francis, New York.
- Sykes I. (2025), “From ‘Girlboss’ to #stayathomegirlfriend”, *European Journal of Cultural Studies*, 28(3), 830–848.
- Terranova T. (2024), *Dopo Internet: Le reti digitali tra capitale e comune*, Nero, Roma.
- Van Dijck J., Poell T., de Waal M. (2018), *The Platform Society*, Oxford University Press, Oxford.
- Vagni T. (2017), “Food (is not) Porn”, *Mediascapes Journal*, (9), 56–69.
- Vizcaíno-Verdú A., Abidin C. (2022), “Music challenge memes on TikTok”, *International Journal of Communication*, 16(26) 131-150.
- Zuboff S. (2019), *Il capitalismo della sorveglianza*, Luiss University Press, Roma.

Le autrici e gli autori

Enrica Amaturò è una sociologa italiana, già professoressa ordinaria di Sociologia generale presso il Dipartimento di Scienze Sociali dell'Università degli Studi di Napoli "Federico II", dove ha insegnato Metodologia della ricerca sociale ed Epistemologia critica delle scienze sociali. È stata Presidente dell'Associazione Italiana di Sociologia (AIS). Le sue ricerche sul digitale analizzano le implicazioni epistemologiche, interrogando il ruolo di dati e algoritmi nella costruzione della conoscenza, nei processi decisionali e nelle nuove configurazioni del potere sociale.

E-mail: enrica.amaturò@unina.it

Caterina Ambrosio è dottoranda in Scienze Sociali e Statistiche presso l'Università degli Studi di Napoli Federico II. La borsa di dottorato è finanziata all'interno del progetto FOSSR - Fostering Open Science in Social Science Research. Tra i principali interessi di ricerca: metodologia della ricerca sociale, digital methods, sociologia digitale, digital awareness. Tra le sue ultime pubblicazioni: *Artificial social research? AI's capabilities and risks in predicting values and attitudes* (con C.C. De Falco, D. Trezza, 2025); *Digital twins: Reconstructing expert narratives on emerging technology*, con R. Di Lorenzo (Rivista di Digital Politics, 2025).

E-mail: caterina.ambrosio@unina.it

Davide Bennato è professore associato di Sociologia dei processi culturali e comunicativi presso il Dipartimento di Scienze Umanistiche dell'Università di Catania. È Presidente del corso di laurea in Scienze e lingue per la comunicazione, presso la medesima Università, e autore del blog *Tecnoetica.it* (<http://tecnoetica.it>). Tra le pubblicazioni più recenti: *La società del XXI secolo. Persone, dati, tecnologie* (Laterza, 2024); *Amanti sintetici. Sesso, relazioni e intimità nell'epoca dell'intelligenza artificiale* (Il Pensiero Scientifico Editore, 2025).

E-mail: davide.bennato@unict.it

Claudia Cantale è sociologa dei processi culturali e comunicativi e docente di sociologia dei media di massa e digitali presso l'Università di Catania. L'area di indagine principale è relativa alle conseguenze del capitalismo, delle piattaforme sulle pratiche delle comunità digitali e delle culture giovanili digitali. Tra le sue pubblicazioni: *Gli Uomini e i testi* (ledizioni.it, 2019) e *Romance e Social Reading. Comunità in rete tra piattaforme e algoritmi* (Carocci, 2024).
E-mail: claudia.cantale@unict.it

Ciro Clemente De Falco è ricercatore in Sociologia generale presso l'Università degli Studi di Napoli "Federico II". Dal 2021 è componente dell'Osservatorio Fast (Futuro, Algoritmi, Società, Tecnologia). Dal 2023 è membro del team di ricerca della Fondazione FAIR (Future Artificial Intelligence Research). Le sue ricerche sul tema degli algoritmi e dell'intelligenza artificiale si concentrano sui processi di mediazione algoritmica, sulle loro implicazioni sociali e sulle trasformazioni del potere e delle disuguaglianze nelle società digitali.
E-mail: ciroclemente.defalco@unina.it

Maria Dentale è ricercatrice in Sociologia generale presso il Dipartimento di Scienze Umane e Sociali Internazionali dell'Università per Stranieri di Perugia. È membro componente del CUG di Ateneo e segretaria della Commissione Paritetica Docenti-Studenti. Tra le sue pubblicazioni più recenti: *Condizionalità e discrezionalità nelle politiche sociali. Una ricerca sulla valutazione delle misure di contrasto alla povertà* (FrancoAngeli, 2022); *La Valutazione d'impatto sociale della didattica digitale dopo il Covid-19. Una prospettiva di ricerca multi-obiettivo e multi-stakeholders* (con V. Lo Presti, a cura di) (FrancoAngeli, 2025).
E-mail: maria.dentale@unistrapg.it

Fiorenza Deriu è professoressa ordinaria di Sociologia generale presso il Dipartimento di Scienze statistiche della Sapienza Università di Roma, dove insegna Sociologia generale, Social network analysis e conduce un Laboratorio di ricerca sociale. È presidente del corso di studi in Statistica, economia e società e dirige un master di II livello in Big data. Metodi statistici per la società della conoscenza. È referente di un centro di competenza su "Reti e testi" per la promozione di un approccio integrato nell'utilizzo di tecniche di Network analysis e di Text mining.
E-mail: fiorenza.deriu@uniroma1.it

Matteo Gerli è professore associato di Sociologia dei processi culturali e comunicativi presso il Dipartimento di Scienze Umane e Sociali Internazionali dell'Università per Stranieri di Perugia, dove svolge le proprie attività di ricerca e insegnamento nell'ambito della sociologia della cultura e dei media. I suoi interessi di ricerca si collocano all'intersezione tra politica, cultura, comunicazione e media, con particolare attenzione ai processi di produzione, legittimazione e circolazione dei saperi e dei beni simbolico-culturali nello spazio pubblico. È autore di due monografie: *Bourdieu, La-tour e la sociologia del giornalismo* (Orthotes, 2024) e *L'Europa della conoscenza* (FrancoAngeli, 2022) – e di numerosi articoli su riviste *peer-reviewed* di fascia A.

E-mail: matteo.gerli@unistrapg.it

Edmondo Grassi è ricercatore in Sociologia generale presso l'Università San Raffaele Roma. I suoi interessi di ricerca si concentrano su Science and Technology Studies (STS), impatto sociale dell'intelligenza artificiale, immaginario, teoria sociologica e pensiero della complessità. È direttore scientifico della European Fairy Tale Route per il Consiglio d'Europa, coordina il GT12 “Récits, fiction, culture et société” dell' AISLF ed è attivo in diverse sezioni dell' AIS. Collabora con numerose riviste scientifiche (tra cui *Humanities & Social Sciences Communications*, *Nature*) ed è co-direttore di *Sociologie*. Nel 2024 ha ricevuto il Premio AIS Giovani come miglior giovane ricercatore per l'articolo *Affective Intersections: Exploring the Ontogenesis of Human-Algorithm Relationships*. Ha pubblicato numerosi articoli e due monografie: *Per una sociologia algomorfica* (FrancoAngeli, 2024) ed *Etica e intelligenza artificiale. Questioni aperte* (Aracne, 2020).

E-mail: edmondo.grassi@uniroma5.it

Rolando Marini è Professore ordinario di Sociologia dei processi culturali e comunicativi presso il Dipartimento di Scienze Umane e Sociali Internazionali dell'Università per Stranieri di Perugia. I suoi interessi di studio e di ricerca si concentrano sulle teorie del rapporto tra media e società, sul giornalismo e sulle dinamiche dello spazio pubblico, con particolare attenzione alle trasformazioni co-generate dall'interazione tra mezzi di informazione, politica e cultura. È autore di diverse monografie, tra cui *Immigrazione e sicurezza* (con S. Tusini, M. Gerli e A. Marcaccio, 2019) e *Mass media e discussione pubblica. Le teorie dell'agenda setting* (2006, Roma-Bari, Laterza), nonché di numerosi articoli pubblicati su riviste *peer-reviewed* di fascia A.

E-mail: rolando.marini@unistrapg.it

Stefania Tusini è professoressa ordinaria di Sociologia generale presso il Dipartimento di Scienze umane e sociali internazionali dell'Università per Stranieri di Perugia. È Presidente del CUG di Ateneo e Delegata rettoriale per le politiche di inclusione. Tra le pubblicazioni più recenti: *Sulla differenza: genere, razza, classe* (con S. Sgavicchia, A. Simoncini, a cura di) (Meltemi Editore, 2025); *Women and Peace Thesis: sicurezza di genere, pace e comportamento degli Stati* (in M. Morazzoni, L. Mundula, a cura di) (CLUEB, 2026).

E-mail: stefania.tusini@unistrapg.it

Questo 
LIBRO

 ti è piaciuto?

Comunicaci il tuo giudizio su:
www.francoangeli.it/opinione



**VUOI RICEVERE GLI AGGIORNAMENTI
SULLE NOSTRE NOVITÀ
NELLE AREE CHE TI INTERESSANO?**



ISCRIVITI ALLE NOSTRE NEWSLETTER

SEGUICI SU:



FrancoAngeli

La passione per le conoscenze

Vi aspettiamo su:

www.francoangeli.it

per scaricare (gratuitamente) i cataloghi delle nostre pubblicazioni

DIVISI PER ARGOMENTI E CENTINAIA DI VOCI: PER FACILITARE
LE VOSTRE RICERCHE.



Management, finanza,
marketing, operations, HR

Psicologia e psicoterapia:
teorie e tecniche

Didattica, scienze
della formazione

Economia,
economia aziendale

Sociologia

Antropologia

Comunicazione e media

Medicina, sanità



Architettura, design,
arte, territorio

Informatica, ingegneria
Scienze

Filosofia, letteratura,
linguistica, storia

Politica, diritto

Psicologia, benessere,
autoaiuto

Efficacia personale

Politiche
e servizi sociali



FrancoAngeli

La passione per le conoscenze

I dati e gli algoritmi rappresentano, oggi, infrastrutture centrali della vita sociale e politica, intervenendo in modo decisivo nei processi che regolano decisioni, opportunità e pratiche quotidiane. Così facendo, orientano l'accesso a risorse e ridefiniscono rapporti di potere, forme di conoscenza e modalità di governo.

Lungi dall'essere strumenti neutrali, essi incorporano visioni del mondo e gerarchie di valore che, attraverso bias e criteri di classificazione opachi, possono rafforzare diseguaglianze sociali esistenti o produrne di nuove. Il volume analizza criticamente questi processi, interrogando il ruolo dei sistemi algoritmici nella configurazione delle diseguaglianze contemporanee e mettendo a fuoco gli spazi di negoziazione, conflitto e resistenza che si aprono attorno al loro utilizzo. Attraverso prospettive teoriche e ricerche empiriche, i contributi qui raccolti mostrano come la dataficazione attraverso ambiti eterogenei (dal welfare alla valutazione delle politiche pubbliche, dai dispositivi di intelligenza artificiale conversazionale ai processi di produzione e validazione della conoscenza scientifica), generando bias, asimmetrie ed effetti differenziati su individui e gruppi sociali. Accanto alle dinamiche di dominio e opacità, un'attenzione particolare è riservata alle pratiche di contestazione, riappropriazione e uso critico delle tecnologie basate sui dati e sugli algoritmi, nonché alle forme di resistenza individuale e collettiva che si sviluppano in risposta ai meccanismi di controllo, classificazione e decisione algoritmica. Nel suo insieme, il testo propone una lettura problematizzante del rapporto tra innovazione tecnologica e società, offrendo strumenti analitici per comprendere la natura ambivalente dei sistemi *data-intensive*, capaci tanto di rafforzare assetti di potere quanto di aprire spazi di critica e resistenza.

Stefania Tusini è professoressa ordinaria di Sociologia generale presso il Dipartimento di Scienze Umane e Sociali Internazionali dell'Università per Stranieri di Perugia.

Davide Bennato è professore associato di Sociologia dei processi culturali e comunicativi presso il Dipartimento di Scienze Umanistiche dell'Università di Catania.

Maria Dentale è ricercatrice in Sociologia generale presso il Dipartimento di Scienze Umane e Sociali Internazionali dell'Università per Stranieri di Perugia.