



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:

<https://www.francoangeli.it/autori/21>

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.

Stereotipi, discriminazioni e disuguaglianze di genere

Una lettura multidisciplinare

a cura di

Stefania Tusini e Maria Dentale



Sociologia

FrancoAngeli 

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: 9788835186045

Isbn e-book Open Access: 9788835195511

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

Pubblicato 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.

Stampa: Global Print s.r.l., Via degli Abeti n. 17/1, 20064 Gorgonzola (MI).

Indice

Introduzione , di <i>Stefania Tusini, Maria Dentale</i>	pag.	9
1. Stereotipi, discriminazioni e disuguaglianze di genere: un quadro analitico	»	9
2. L'articolazione del volume	»	13
Riferimenti bibliografici	»	17
 1. <i>Gender data bias</i> in una prospettiva sociologica , di <i>Stefania Tusini, Maria Dentale</i>	»	19
1. Introduzione	»	19
2. <i>Gender data bias</i> , big data e processi di <i>machine learning</i>	»	20
3. Dimensione intersezionale del bias e bias cognitivo	»	23
4. <i>Gender data gap</i> e algoritmi come dispositivi di potere	»	26
5. Disuguaglianze prodotte dal <i>gender data bias</i>	»	30
6. Conclusioni	»	37
Riferimenti bibliografici	»	39
 2. Le streghe. Note storiche su alterità e pensiero magico , di <i>Valerio De Cesaris</i>	»	44
1. Introduzione	»	44
2. Caccia alle streghe	»	50
3. Il linguaggio magico e la confessione	»	52

4. La strega, l'ebreo vampiro e il sangue dei cristiani	pag.	56
5. Uno dei più grandi errori giudiziari nella storia dell'Occidente	»	59
Riferimenti bibliografici	»	60
3. Catturare il genere. Note su femminismo neoliberale, diversity management e gender mainstreaming , di <i>Ouns Mornagui, Alessandro Simoncini</i>	»	63
1. Neoliberalismo mutante, processi di neoliberalizzazione e femminismo neoliberale	»	63
2. Governare la differenza. Femminismo istituzionale e managerializzazione del genere	»	66
3. Integrare il genere. Promesse e neutralizzazioni del mainstreaming	»	69
4. Conclusioni	»	71
Riferimenti bibliografici	»	72
4. Parità di genere e lavoro sportivo , di <i>Alberto Azara</i>	»	76
1. La disapplicazione del principio di parità di genere nel settore sportivo	»	76
2. Le origini del problema	»	80
3. La riforma del lavoro sportivo	»	83
4. Segue. Alcuni rilievi critici	»	84
Riferimenti bibliografici	»	87
5. Donne e diplomazia nell'Italia repubblicana: un itinerario storico , di <i>Federico Niglia, Lucrezia Luci</i>	»	89
1. Cenni storici sull'esclusione delle donne dalla diplomazia	»	89
2. Le battaglie delle pioniere	»	92
3. La Farnesina apre le porte: vecchi impedimenti e nuovi traguardi	»	95
4. Donne in diplomazia o diplomazia delle donne: conclusioni (provvisorie) su un percorso	»	100
Riferimenti bibliografici	»	102

6. Identità di genere e pratiche di autonarrazione: un'indagine qualitativa condotta su un campione di studentesse internazionali dell'Università per Stranieri di Perugia , di <i>Maria Flavia Maiorano, Siriana Sgavicchia</i>	pag.	105
1. "Fuori tema". Identità di genere e narrazione	»	105
2. Impostazione metodologica del questionario	»	108
3. Le risposte al questionario	»	110
4. Conclusioni	»	117
Riferimenti bibliografici	»	119
 7. La parità di genere nello sport: gli artt. 39 e 40 del d.lgs. 36/2021 come strumenti di inclusione sociale , di <i>Luigi Gradante, Alberto Azara</i>	»	120
1. Verso un diritto sportivo inclusivo	»	120
2. Oltre il dilettantismo imposto: la riforma del lavoro sportivo e la parità di genere	»	121
3. L'art. 39 del d.lgs. n. 36/2021: un passo avanti per il professionismo sportivo femminile	»	124
4. La promozione della parità di genere nello sport: l'art. 40 del d.lgs. n. 36/2021	»	127
Riferimenti bibliografici	»	133
 8. Femminicidi in Africa: retaggi coloniali e violenze contemporanee , di <i>Federica Guazzini</i>	»	134
1. Introduzione: un fenomeno globale, una storia plurale	»	134
2. Nominare e misurare: la questione della produzione del dato	»	136
3. Una cornice integrata: patriarcato, ecologia della violenza e intersezionalità	»	137
4. Genealogie coloniali della disuguaglianza e della violenza	»	139
5. Violenza contemporanea: persistenze e trasformazioni	»	143
6. Dalla norma alla prevenzione: diritti, istituzioni e femminismi africani	»	148

7. Conclusioni	pag.	151
Riferimenti bibliografici	»	152
9. L'8 marzo come posta in gioco nell'arena mediale. Continuità e trasformazioni nella rappresentazione delle questioni di genere, di <i>Rolando Marini, Matteo Gerli</i>		
	»	158
1. Introduzione	»	158
2. Metodologia dello studio: <i>content analysis</i> degli articoli connessi alla costruzione dei contenuti mediatici dell'8 marzo	»	162
3. Il tema del <i>genere</i> come spazio relazionale, tra notiziabilità e strategie narrative	»	165
4. Due 8 marzo a confronto, tra celebrazione e controversia simbolica	»	169
5. Conclusioni	»	177
Riferimenti bibliografici	»	181
10. Il paradosso di genere nelle scienze e nella governance globale dell'acqua: verso un paradigma trasformativo, di <i>Chiara Biscarini, Maura Marchegiani, Gaia Proietti</i>		
	»	185
1. Introduzione: l'acqua come frontiera della complessità	»	185
2. Il quadro giuridico internazionale: tra universalità formale ed effettività problematica	»	187
3. L'evidenza empirica della sottorappresentazione istituzionale	»	190
4. Conclusioni: verso una <i>governance</i> idrica <i>gender-transformative</i> tra diplomazia scientifica, inclusione e sicurezza idrica	»	197
Riferimenti bibliografici	»	199
Gli autori e le autrici	»	203

1. Gender data bias e disuguaglianze in una prospettiva sociologica

di *Stefania Tusini, Maria Dentale**

1. Introduzione

Negli ultimi decenni, la crescente disponibilità di dati digitali e lo sviluppo di tecniche avanzate di analisi computazionale hanno profondamente trasformato il modo in cui i fenomeni sociali vengono osservati, interpretati e governati. Questo processo, comunemente definito come *datafication* della società, si fonda sull'assunto implicito che la realtà sociale possa essere tradotta in dati "misurabili", analizzabili e impiegabili per orientare decisioni pubbliche e private. Tuttavia, tale presupposto richiama una concezione positivista dei dati come entità neutrali e oggettive, prospettiva ampiamente problematizzata dalla sociologia della conoscenza, dalla sociologia della scienza e dagli studi sulla quantificazione (Latour e Woolgar, 1979; Latour, 1987; Porter, 1995; Desrosières, 1993). Come affermano Berger e Luckmann, «la realtà è socialmente costruita» (1966, p. 13); concetto ribadito tra gli altri da Bourdieu che sottolinea che «l'oggetto scientifico non è dato, ma costruito» (Bourdieu *et al.*, 1968, p. 51). In questa prospettiva, i dati non costituiscono semplici rappresentazioni passive della realtà, bensì prodotti sociali situati. Anche la letteratura contemporanea conferma questa impostazione: «big data are not a neutral representation of reality» (Boyd, Crawford, 2012, p. 663), mentre Gitelman evidenzia come parlare di «raw data is an oxymoron» (2013, p. 1).

* Il capitolo è il risultato della riflessione congiunta delle autrici.

In questo quadro teorico, il *gender data bias* rappresenta un caso paradigmatico di come i dati possano riflettere e rafforzare le disuguaglianze sociali di genere. Il bias emerge lungo l'intero ciclo di vita dei dati e non può essere ridotto a un problema tecnico. Come mostrano studi recenti nel campo sanitario, ad esempio, i dataset osservazionali risultano spesso incompleti e selettivi: «observational datasets [...] may not necessarily include all the attributes relevant for the analysis» (Díaz-Rodríguez *et al.*, 2023, pp. 1-2). Questa incompletezza non è casuale, ma riflette priorità istituzionali e limiti strutturali che influenzano la rappresentazione della realtà.

2. Gender data bias, big data e processi di machine learning

Nel contesto dei big data, tali dinamiche risultano ulteriormente amplificate. I dati generati dalle piattaforme digitali sono il prodotto di processi sociali caratterizzati da disuguaglianze di accesso e partecipazione. Nel caso di OpenStreetMap¹, ad esempio, la produzione dei dati è fortemente sbilanciata: «volumes of overall activity [...] remain significantly dominated by men» (Gardner *et al.*, 2020, p. 1605; Lam *et al.*, 2011). La piattaforma è un esempio perfetto di produzione sociale dei dati in cui i contributori non sono rappresentativi della popolazione e la partecipazione è fortemente sbilanciata per genere. Infatti, come mostrano gli studi su citati, la maggioranza dei contributori è maschile e di conseguenza alcune informazioni (ad esempio sui servizi legati alla cura) sono meno mappate. Ciò produce un bias nei dati geografici e implica che i dataset risultanti riflettono una prospettiva parziale, che non rappresenta adeguatamente l'intera popolazione.

La produzione collaborativa dei dati geografici pertanto non è neutrale, ma riflette la composizione sociale dei contributori con la

¹ OpenStreetMap è una piattaforma collaborativa di mappatura *open source*, basata sul contributo volontario degli utenti, che consente la produzione condivisa di dati geografici (<https://www.openstreetmap.org/#map=5/42.09/12.56>). La sua natura partecipativa la rende un caso paradigmatico per l'analisi delle disuguaglianze nella produzione dei dati.

conseguenza che «what gets mapped reflects the priorities of those who do the mapping» (Stephens, 2013, p. 5). Di conseguenza, tali dati «fail to embody the [...] interests of the wider community» (Gardner *et al.*, 2020, p. 1604), evidenziando come il bias sia già incorporato nella fase di produzione.

Gli algoritmi di *machine learning* contribuiscono a rafforzare queste distorsioni. Poiché apprendono dai dati, essi tendono a riprodurre le disuguaglianze esistenti. Come osserva O'Neil, «models are opinions embedded in mathematics» (2016, p. 21), sottolineando come i modelli non siano neutrali ma incorporino scelte e valori.

Nel campo della *computer vision*, ad esempio, numerosi studi empirici hanno evidenziato la presenza di bias sistematici nei sistemi di riconoscimento e classificazione delle immagini, in particolare in relazione al genere (Scheuerman, 2019). Come mostrano Schwemmer *et al.*, «women in images are recognized at substantially lower rates» (2020, p. 2), indicando una minore precisione degli algoritmi nel riconoscere i volti e le caratteristiche delle donne rispetto a quelli degli uomini. Questo fenomeno è strettamente legato alla composizione dei dataset di addestramento, spesso sbilanciati e caratterizzati da una rappresentazione meno varia delle soggettività femminili.

Oltre alla dimensione quantitativa dell'accuratezza, emerge una dimensione qualitativa altrettanto rilevante. Le immagini che ritraggono donne tendono infatti a essere associate più frequentemente a caratteristiche legate all'aspetto fisico, mentre le immagini maschili sono più spesso descritte in termini di ruolo, attività o status sociale. Come evidenziano gli studi, le annotazioni riferite alle donne risultano significativamente più concentrate su elementi estetici, contribuendo a una rappresentazione stereotipata e riduttiva (Schwemmer *et al.*, 2020). Questo tipo di classificazione non è neutrale, ma riflette schemi culturali preesistenti che vengono incorporati nei sistemi tecnologici.

Questa differenza non può essere interpretata come una semplice imperfezione tecnica, ma deve essere letta alla luce delle strutture simboliche che organizzano la percezione sociale. In termini *bourdieusiani*, tali classificazioni algoritmiche contribuiscono a una forma di violenza simbolica, intesa come quel processo attraverso cui le gerarchie sociali vengono interiorizzate e naturalizzate, apparendo come ovvie e legittime (Bourdieu, 1979). Gli algoritmi, apprendendo

da dati che riflettono rappresentazioni culturali sedimentate, finiscono per incorporare e riprodurre tali schemi, trasformandoli in output apparentemente neutrali. In questo modo, la distinzione tra uomini e donne viene codificata e reiterata attraverso categorie che, come detto, associano sistematicamente le donne alla dimensione estetica e gli uomini a quella funzionale o produttiva.

Questi risultati assumono una particolare rilevanza se interpretati alla luce del funzionamento dei modelli di apprendimento automatico. Gli algoritmi di *machine learning*, infatti, non fanno altro che apprendere pattern dai dati disponibili; di conseguenza, se i dataset contengono rappresentazioni sbilanciate o stereotipate, tali caratteristiche vengono non solo replicate, ma spesso amplificate nei risultati. In questo senso, il bias non è un errore accidentale, ma una proprietà emergente del sistema sociotecnico nel suo complesso. Anche O'Neil (2016) lo sottolinea, evidenziando come le scelte implicite nei dati e nei modelli si traducano in esiti sistematici.

Il carattere particolarmente insidioso di questo processo risiede nel fatto che la violenza simbolica operata dai sistemi di intelligenza artificiale è mediata da tecnologie percepite come oggettive e scientificamente fondate, ma tale dimensione normativa tende a rimanere invisibile agli utenti. Di conseguenza, le classificazioni algoritmiche non solo riflettono stereotipi preesistenti, ma contribuiscono a rafforzarli e legittimarli, rendendo più difficile la loro contestazione. In questo senso, la *computer vision* non si limita a rappresentare la realtà sociale, ma partecipa attivamente alla sua costruzione, operando come dispositivo contemporaneo di riproduzione delle disuguaglianze simboliche.

La conseguenza è che i sistemi di *computer vision* contribuiscono alla riproduzione di stereotipi di genere, rafforzando rappresentazioni sociali esistenti. Questo processo di amplificazione tecnologica degli stereotipi è particolarmente problematico in contesti applicativi come la sicurezza, il marketing o i social media, dove le classificazioni algoritmiche influenzano direttamente la visibilità, l'accesso alle risorse e le opportunità degli individui.

In definitiva, il caso della *computer vision* dimostra come il *gender data bias* non si limiti a riflettere le disuguaglianze esistenti, ma contribuisca attivamente a consolidarle, trasformando rappresentazioni socialmente costruite in output tecnologici percepiti come og-

gettivi. Ciò evidenzia la necessità di un approccio critico alla progettazione dei sistemi di intelligenza artificiale, che tenga conto non solo della performance tecnica, ma anche delle implicazioni sociali e culturali dei modelli utilizzati.

3. Dimensione intersezionale del bias e bias cognitivo

Un elemento particolarmente rilevante di questo ragionamento riguarda la dimensione intersezionale del bias. Le discriminazioni algoritmiche, infatti, non colpiscono tutti i gruppi sociali nello stesso modo, ma tendono ad amplificarsi quando si sovrappongono più fattori di vulnerabilità, come genere, razza, colore della pelle, classe sociale o appartenenza geografica. In questo senso, il bias non può essere interpretato come un fenomeno lineare o “a variabile unica”, ma come il prodotto di relazioni sociali stratificate che si riflettono nei dati e nei sistemi di intelligenza artificiale.

Uno degli esempi più evidenti emerge dallo studio di Yoder-Himes *et al.* (2022) sui software di *proctoring* automatico utilizzati nelle università durante gli esami online. Gli autori mostrano che gli algoritmi di riconoscimento facciale tendevano a segnalare più frequentemente gli studenti con pelle scura come “sospetti”, ma soprattutto che «women with the darkest skin tones were far more likely [...] to be flagged» (ivi, p. 1). Ciò significa che il sistema non discriminava semplicemente sulla base del genere o della razza separatamente, bensì attraverso la loro combinazione: le donne nere risultavano il gruppo maggiormente penalizzato.

Questo tipo di evidenza conferma la prospettiva dell’intersezionalità sviluppata da Crenshaw (1989, 1991), secondo cui le forme di disuguaglianza non agiscono isolatamente, ma si intrecciano producendo effetti cumulativi. Nei sistemi algoritmici, tali effetti derivano spesso dai dataset di addestramento, storicamente sbilanciati verso soggetti maschi, bianchi e occidentali, e dalla riproduzione automatizzata di stereotipi presenti nella società. Di conseguenza, gruppi già marginalizzati rischiano di essere ulteriormente invisibilizzati o penalizzati.

La dimensione intersezionale del bias è stata osservata anche in altri ambiti. Ad esempio, Zheng *et al.* (2023) evidenziano che Wikipedia tende a citare meno frequentemente lavori scientifici prodotti

da donne e da autori affiliati a paesi non anglofoni, con effetti particolarmente forti nei settori STEM più matematici. Anche in questo caso, genere e provenienza geografica si combinano producendo svantaggi differenziati.

L'intersezionalità del bias assume particolare rilievo anche nel settore sanitario e psicologico. Rodriguez-Seijas, Morgan e Zimmerman (2024) mostrano che i pazienti transgender e *gender diverse* ricevono diagnosi di disturbo borderline significativamente più frequentemente rispetto ai pazienti *cisgender*, anche controllando la presenza di sintomi specifici. Gli autori suggeriscono che stereotipi e aspettative sociali influenzino implicitamente le valutazioni cliniche, dimostrando come identità di genere e stigma sociale possano interagire nella produzione di bias diagnostici.

Questi studi evidenziano una volta di più che l'iniquità algoritmica non dipende solo da errori tecnici, ma riflette strutture sociali più profonde. Gli algoritmi apprendono dai dati storici e, se tali dati incorporano disuguaglianze pregresse, il rischio è quello di automatizzare e rafforzare discriminazioni già esistenti. Per questo motivo, numerosi autori sostengono la necessità di adottare approcci di *causal fairness*² e di *auditing* algoritmico³, capaci di considerare simultaneamente variabili sociali differenti e le loro interazioni (Díaz-Rodríguez *et al.*, 2023).

Un ulteriore livello di analisi riguarda il rapporto tra bias cognitivo⁴ e bias nei dati. I processi di raccolta, classificazione e interpreta-

² Gli approcci di *causal fairness* mirano a distinguere gli effetti direttamente attribuibili a caratteristiche sensibili (come genere o razza) da quelli derivanti da fattori sociali, economici o strutturali che agiscono come mediatori o variabili confondenti, evitando così che gli algoritmi riproducano discriminazioni pregresse (Pearl, 2009; Díaz-Rodríguez *et al.*, 2023).

³ L'*auditing* algoritmico è l'insieme delle procedure utilizzate per analizzare, verificare e valutare il funzionamento di un algoritmo o di un sistema di intelligenza artificiale, con l'obiettivo di identificare eventuali problemi di bias e discriminazioni, errori sistematici, mancanza di trasparenza, violazioni etiche o normative. Questa procedura in pratica è una sorta di revisione o ispezione degli algoritmi e serve a capire come un algoritmo prende decisioni, quali dati utilizza, se produce effetti discriminatori su determinati gruppi sociali e se i risultati sono affidabili, equi e spiegabili.

⁴ Per bias cognitivo si intende una distorsione sistematica del processo di percezione, interpretazione o giudizio della realtà, prodotta dal modo in cui il cervello

zione dell'informazione non sono infatti neutrali, come detto, ma risultano profondamente influenzati da schemi cognitivi, stereotipi culturali e aspettative socialmente condivise che orientano ciò che viene osservato, registrato e considerato rilevante. Le decisioni relative a quali dati raccogliere, quali categorie utilizzare, quali variabili includere o escludere e quali fenomeni ritenere “normali” riflettono inevitabilmente visioni del mondo storicamente situate e rapporti di potere presenti nella società.

In questo senso, il *gender data bias* può essere interpretato come il risultato dell'interazione tra bias cognitivi individuali, norme sociali e strutture istituzionali. I bias cognitivi agiscono a livello micro, influenzando la percezione e il giudizio dei singoli individui: ad esempio, stereotipi impliciti possono portare ricercatori, clinici o sviluppatori a considerare il corpo maschile come standard universale o a interpretare determinati comportamenti femminili secondo categorie stereotipate. A livello meso e macro, invece, intervengono norme organizzative, pratiche professionali e assetti istituzionali che consolidano tali rappresentazioni trasformandole in procedure standardizzate, classificazioni statistiche e modelli algoritmici.

Questo meccanismo è particolarmente evidente nei sistemi di intelligenza artificiale e *machine learning*, che apprendono da dati storici prodotti all'interno di contesti sociali già caratterizzati da disuguaglianze. Gli algoritmi non generano autonomamente il bias, ma tendono piuttosto a incorporare e riprodurre le distorsioni presenti nei dati di addestramento. Come osserva Noble (2018), i sistemi algorit-

umano elabora le informazioni. I bias cognitivi influenzano il modo in cui gli individui selezionano, ricordano e valutano dati ed eventi, portando spesso a decisioni non pienamente razionali o oggettive. Questi meccanismi agiscono in modo inconsapevole e rappresentano scorciatoie mentali (*heuristics*) che permettono di elaborare rapidamente grandi quantità di informazioni. Tuttavia, proprio perché semplificano la realtà, possono produrre stereotipi, generalizzazioni e valutazioni distorte (Tversky e Kahneman, 1974; Kahneman, 2011). Tversky e Kahneman (1974) individuano anche molti altri tipi di bias, tra cui: *availability bias* (quando si tende a valutare la probabilità di un evento in base alla facilità con cui ci vengono in mente esempi e pertanto si sovrastimano eventi molto visibili); *representativeness heuristic* (quando si giudica qualcosa in base a quanto assomiglia a uno stereotipo o modello mentale, ignorando spesso le probabilità reali); *anchoring bias* (quando si tende a farsi influenzare dal primo dato ricevuto, che fa da ancora).

mici possono funzionare come strumenti di amplificazione delle gerarchie sociali esistenti, poiché automatizzano categorie e associazioni costruite storicamente. Analogamente, Friedman e Nissenbaum (1996) hanno evidenziato che il bias nei sistemi informatici può derivare non solo da errori tecnici, ma anche dai valori e dalle assunzioni implicite incorporati nei processi di progettazione tecnologica.

Anche le piattaforme digitali e i sistemi di classificazione automatizzata, come detto, mostrano chiaramente questa relazione tra dimensione cognitiva e produzione dei dati (Schwemmer *et al.*, 2020). Questo tipo di classificazione non riflette semplicemente differenze “oggettive”, ma riproduce stereotipi culturali sedimentati nei dataset e nei modelli cognitivi dei progettisti.

Il problema riguarda dunque non solo la qualità tecnica dei dati, ma anche i processi sociali che ne guidano la produzione. Le categorie statistiche e gli standard informativi utilizzati nelle banche dati non sono neutri: definiscono quali soggetti risultano visibili, quali esperienze vengono considerate rilevanti e quali invece rimangono marginali o invisibili. Di conseguenza, il *gender data bias* non rappresenta esclusivamente una distorsione metodologica, ma costituisce l'espressione di rapporti sociali e culturali che si sedimentano nei dati e vengono successivamente riprodotti dalle tecnologie digitali.

4. Gender data gap e algoritmi come dispositivi di potere

Il fenomeno in esame è inoltre strettamente connesso al *gender data gap*, ovvero alla mancanza sistematica di dati relativi alle donne. Ciò che non viene rilevato tende infatti a diventare invisibile, con conseguenze significative in termini di rappresentazione sociale, definizione delle priorità e allocazione delle risorse pubbliche. L'assenza di dati costituisce quindi una forma di bias per omissione che si affianca al bias per distorsione, poiché non solo i dati possono essere raccolti in modo sbilanciato, ma alcuni fenomeni possono addirittura non essere registrati o studiati (Criado Perez, 2019)⁵.

⁵ Criado Perez nel corso della sua analisi individua diverse forme ricorrenti di distorsione legate alla produzione, raccolta e utilizzo dei dati. Oltre al *bias per omissione*

Un esempio particolarmente noto riguarda la medicina cardiovascolare. Per molti decenni gli studi clinici sull'infarto sono stati condotti prevalentemente su campioni maschili, contribuendo a definire come “tipici” sintomi maggiormente presenti negli uomini, come il forte dolore al petto irradiato al braccio sinistro. Di conseguenza, i sintomi più frequenti nelle donne – nausea, affaticamento, dolore alla schiena o difficoltà respiratorie – sono stati a lungo sottovalutati o considerati atipici, aumentando il rischio di diagnosi tardive o errate (Holdcroft, 2007).

Un altro esempio riguarda la sicurezza automobilistica. I crash test utilizzati per decenni dall'industria automobilistica sono stati costruiti principalmente su manichini modellati sul corpo maschile medio adulto. L'assenza di dati adeguati sulle caratteristiche fisiche femminili ha contribuito a una minore protezione delle donne negli incidenti stradali: alcune ricerche mostrano infatti che le donne hanno una probabilità significativamente maggiore di subire lesioni gravi in incidenti automobilistici comparabili (Criado Perez, 2019).

Anche nel mercato del lavoro il *gender data gap* produce effetti rilevanti. In molti paesi il lavoro domestico e di cura svolto prevalentemente dalle donne non viene completamente contabilizzato nelle statistiche economiche tradizionali, rendendo invisibile una parte es-

(assenza di dati sulle donne) e al *bias per distorsione* (dati raccolti o interpretati in modo sbilanciato), emergono almeno altre cinque forme rilevanti di bias: *bias androcentrico* (considerare implicitamente il corpo, l'esperienza e il comportamento maschile come standard universale; ad esempio crash test automobilistici basati su manichini maschili); *bias di progettazione* (il modo in cui tecnologie, prodotti, spazi e servizi vengono progettati senza considerare differenze di genere con la conseguenza che il design incorpora implicitamente un utente di tipo maschile; ad esempio smartphone troppo grandi per molte mani femminili); *bias di campionamento* (quando i campioni utilizzati nella ricerca non rappresentano adeguatamente la popolazione in base al genere; ad esempio dati sul lavoro costruiti prevalentemente sul lavoro retribuito maschile); *bias istituzionale* (distorsioni che derivano da procedure, norme e pratiche organizzative consolidate e discriminatorie in base al genere; ad esempio modelli economici che ignorano il lavoro di cura); *bias algoritmico* (quando gli algoritmi apprendono da dati storicamente sbilanciati in termini di genere; ad esempio algoritmi di reclutamento professionale che penalizzano CV femminili); *bias di invisibilità* (forma estrema del bias per omissione: alcuni fenomeni che riguardano prevalentemente le donne non vengono nemmeno concepiti come oggetti di rilevazione; ad esempio mancanza di dati sulla mobilità quotidiana legata al lavoro di cura) (2019, *passim*).

senziale del lavoro sociale necessario al funzionamento dell'economia. Questa invisibilità statistica contribuisce a sottovalutare il peso economico e sociale del lavoro di cura e può influenzare negativamente le politiche di welfare e di sostegno familiare (Folbre, 2006).

Nel contesto digitale e algoritmico, l'assenza di dati può inoltre compromettere il funzionamento dei sistemi di intelligenza artificiale. Ad esempio, software di riconoscimento vocale addestrati prevalentemente su voci maschili tendono a riconoscere con minore accuratezza le voci femminili, mentre sistemi di riconoscimento facciale possono avere performance peggiori su donne con pelle scura a causa della scarsa rappresentazione di questi gruppi nei dataset di addestramento. In questi casi, l'invisibilità nei dati si traduce direttamente in minore accuratezza tecnologica e maggiore rischio di discriminazione (Schwemmer *et al.*, 2020; Yoder-Himes *et al.*, 2022).

Questi esempi mostrano come il *gender data gap* non rappresenti semplicemente una lacuna informativa, ma un problema strutturale che influenza la produzione della conoscenza, lo sviluppo tecnologico e la distribuzione delle risorse, contribuendo alla riproduzione delle disuguaglianze di genere nella società.

Dal punto di vista teorico, queste evidenze confermano la necessità di interpretare i dati non come elementi neutri o puramente tecnici, ma come veri e propri dispositivi di potere. La produzione dei dati, la costruzione delle categorie statistiche e i processi algoritmici attraverso cui le informazioni vengono selezionate, ordinate e interpretate riflettono infatti rapporti sociali, valori culturali e priorità istituzionali. I dati non descrivono semplicemente la realtà: contribuiscono anche a costruirla, stabilendo quali fenomeni meritino attenzione, quali soggetti risultino visibili e quali invece rimangano marginali o esclusi.

In questa prospettiva, il pensiero di Foucault offre un quadro teorico particolarmente rilevante. Secondo Foucault (1975) i sistemi di conoscenza non sono separabili dalle forme di potere che li producono: sapere e potere si rafforzano reciprocamente attraverso pratiche di classificazione, osservazione e controllo. Le istituzioni moderne – dalla medicina alla statistica, dalla scuola al sistema giudiziario – organizzano la realtà attraverso categorie che definiscono ciò che è normale, deviante, rilevante o invisibile. I dati e le tecnologie digitali possono essere interpretati come una continuazione contemporanea di questi meccanismi di disciplinamento e governo delle popolazioni.

Gli algoritmi rappresentano dunque forme contemporanee di potere opaco, che operano attraverso procedure automatizzate spesso difficilmente comprensibili anche dagli stessi utilizzatori. A differenza delle tradizionali forme di potere visibile e centralizzato, il potere algoritmico agisce in modo diffuso e silenzioso, incorporandosi nelle piattaforme digitali, nei sistemi di raccomandazione, nei modelli predittivi e nei processi decisionali automatizzati. Attraverso l'analisi massiva dei dati, questi sistemi influenzano l'accesso alle informazioni, la selezione delle opportunità lavorative, l'erogazione di servizi sanitari, la concessione di credito e persino la sorveglianza sociale.

Diversi autori hanno evidenziato come questa infrastruttura algoritmica rischi di produrre nuove forme di disuguaglianza e soprattutto di controllo sociale. Gillespie (2014), ad esempio, sottolinea che gli algoritmi non essendo strumenti neutrali, influenzano la visibilità delle informazioni e la costruzione della rilevanza sociale. Analogamente, Pasquale (2015) descrive i sistemi algoritmici come *black boxes*, cioè meccanismi opachi il cui funzionamento risulta spesso inaccessibile ai cittadini e difficilmente sottoponibile a controllo democratico. Anche Zuboff (2019) evidenzia come il capitalismo della sorveglianza trasformi i dati personali in risorse economiche e strumenti di previsione comportamentale, rafforzando nuove forme di asimmetria di potere tra piattaforme digitali e individui.

Nel caso del *gender data bias*, questo processo assume implicazioni particolarmente rilevanti. Se i dati riflettono disuguaglianze storiche e stereotipi culturali, gli algoritmi rischiano di naturalizzare tali squilibri, trasformandoli in decisioni apparentemente oggettive e neutrali. In questo modo, discriminazioni socialmente costruite possono essere reinterpretate come semplici "evidenze statistiche", rendendo più difficile riconoscerne la natura politica e strutturale. L'opacità dei sistemi algoritmici contribuisce inoltre a ridurre la possibilità di contestare le decisioni automatizzate, poiché i criteri utilizzati risultano spesso invisibili o incomprensibili per gli individui coinvolti.

Alla luce di quanto discusso, appare evidente che il *gender data bias* non possa essere affrontato esclusivamente attraverso soluzioni tecniche. È necessario sviluppare un approccio critico ai dati, capace di integrare competenze tecniche e riflessione sociologica. Come sottolineano D'Ignazio e Klein, «data [...] are the products of unequal social relations» (2020, p. 4). In questo senso, il superamento del bias

richiede un cambiamento di paradigma nella produzione e nell'uso dei dati.

Le questioni etiche e politiche che emergono riguardano quindi non solo la correttezza tecnica degli algoritmi, ma anche la trasparenza, la responsabilità e il controllo democratico delle tecnologie digitali. Per questo motivo, negli ultimi anni è cresciuta l'attenzione verso pratiche di *auditing* algoritmico (vedi nota 3), *explainable* AI e *governance* etica dei dati, considerate strumenti fondamentali per limitare gli effetti discriminatori e rendere maggiormente visibili i meccanismi di potere incorporati nei sistemi automatizzati.

Come visto, le dinamiche fin qui descritte non si limitano a produrre distorsioni nella rappresentazione della realtà, ma generano effetti concreti nei diversi ambiti della vita sociale, contribuendo alla riproduzione delle disuguaglianze. In questo senso, come accennato, il *gender data bias* deve essere analizzato anche in relazione alle sue conseguenze materiali.

5. Disuguaglianze prodotte dal *gender data bias*

Abbiamo visto come il *gender data bias* non rappresenti soltanto una distorsione nella rappresentazione dei fenomeni sociali, ma costituisca un potente meccanismo di produzione e riproduzione delle disuguaglianze. L'uso crescente di dati e algoritmi nei processi decisionali ha infatti ampliato la portata delle conseguenze del bias, trasformando errori o omissioni nella raccolta dei dati in effetti sistemici che incidono su opportunità, risorse e diritti. In questo senso, come affermato, il bias non si limita a riflettere le disuguaglianze esistenti, ma contribuisce attivamente alla loro stabilizzazione e legittimazione, rendendole meno visibili e più difficili da contestare.

Gli effetti del *gender data bias* si manifestano in vari contesti.

Nel sistema accademico, ad esempio, le donne continuano a risultare penalizzate nei percorsi di carriera, nonostante l'aumento della presenza femminile nell'istruzione superiore e nella ricerca scientifica. Le disuguaglianze emergono in molteplici dimensioni: accesso alle posizioni permanenti, progressione di carriera, retribuzioni, accesso ai ruoli di leadership, riconoscimento scientifico e distribuzione delle opportunità di finanziamento. Bond *et al.*, analizzando i dati

relativi alle facoltà di Farmacia negli Stati Uniti, mostrano che «women [...] had a significantly lower odds ratio of promotion» rispetto ai colleghi uomini (2024, p. 1). Le autrici evidenziano inoltre che le donne presentano minori probabilità di raggiungere posizioni apicali e registrano gap salariali che tendono ad ampliarsi con l'avanzamento della carriera. Studi sulle pratiche di reclutamento accademico mostrano in effetti come criteri apparentemente neutri possano riprodurre disuguaglianze di genere, rendendo invisibili i processi che generano asimmetrie («the conception of organizations [...] as neutral is one of the main causes of the persistence of gender inequalities»; Picardi, 2029, p. 1).

Le disuguaglianze prodotte dal *gender data bias* si manifestano anche nella produzione e nella diffusione della conoscenza. Studi recenti mostrano che le pubblicazioni scientifiche delle donne risultano meno citate rispetto a quelle degli uomini, contribuendo a una distribuzione asimmetrica dell'autorevolezza accademica. Zheng *et al.* evidenziano che «publications by women are cited less [...] than those by men» (2023, p. 220), sottolineando come i sistemi di citazione e visibilità scientifica riflettano e amplifichino le disuguaglianze di genere. Questo fenomeno non solo limita il riconoscimento del lavoro delle ricercatrici, ma influisce anche sulla produzione della conoscenza, orientando l'attenzione verso determinati temi e prospettive a discapito di altri.

La minore visibilità scientifica delle donne produce effetti cumulativi lungo l'intera carriera accademica. Poiché citazioni, impact factor e indicatori bibliometrici rappresentano oggi criteri centrali nei processi di valutazione della ricerca STEM, nei concorsi e nell'assegnazione dei finanziamenti, una minore citabilità si traduce in minori opportunità professionali e istituzionali, creando un corto circuito. In questo modo, il bias nella circolazione del sapere alimenta ulteriormente le disuguaglianze di carriera, generando un circolo cumulativo di svantaggio.

Queste disuguaglianze non dipendono esclusivamente da differenze individuali di produttività o competenze, ma riflettono meccanismi strutturali profondamente radicati nelle istituzioni accademiche. La letteratura sociologica e femminista ha mostrato come reti professionali informali, pratiche di cooptazione, criteri di valutazione apparentemente neutri e modelli organizzativi costruiti storicamente sul

lavoratore maschile contribuiscano alla riproduzione delle disparità di genere. Picardi (2019), ad esempio, parla di *glass door*, evidenziando come i processi di reclutamento e selezione accademica possano produrre nuove forme di segregazione di genere anche in presenza di regole formalmente meritocratiche. Anche strumenti di *assessment* largamente utilizzati, come le valutazioni degli studenti, possono incorporare bias impliciti, portando a una sistematica sottovalutazione delle docenti donne indipendentemente dalla qualità dell'insegnamento (MacNell *et al.*, 2015; Potvin e Hazari, 2016; Valencia, 2022).

Analogamente, come detto, anche la produzione, la diffusione e la legittimazione del sapere scientifico risultano influenzate da bias strutturali. Alcuni autori (Zheng *et al.*, 2023) mostrano che gli articoli prodotti da ricercatrici e da studiosi affiliati a paesi non anglofoni risultano meno visibili e meno frequentemente citati, con effetti particolarmente marcati nei settori STEM a forte componente matematica. Questo fenomeno contribuisce alla costruzione di una gerarchia della conoscenza in cui alcuni soggetti e alcune aree geografiche risultano maggiormente riconosciuti come autorevoli.

Tali dinamiche confermano come il sistema scientifico non costituisca uno spazio completamente neutrale o meritocratico, ma sia attraversato da rapporti di potere, norme culturali e meccanismi di esclusione che influenzano la produzione stessa della conoscenza. Il problema non riguarda soltanto la sottorappresentazione numerica delle donne, ma anche la definizione di ciò che viene considerato scientificamente autorevole, rilevante e degno di riconoscimento pubblico. Questi meccanismi, come detto, non si limitano alla fase di valutazione *ex post*, ma operano lungo l'intero ciclo di produzione del sapere.

Il bias incorporato nei dati e negli algoritmi utilizzati per analizzare la produzione scientifica o supportare decisioni istituzionali può amplificare queste disuguaglianze. Sistemi automatici e modelli predittivi, se costruiti su dati non rappresentativi o incompleti, rischiano di produrre esiti distorti e rafforzare allocazioni discriminatorie di risorse e opportunità. Nel loro insieme, queste evidenze mostrano come il *gender data bias* non sia solo un problema tecnico, ma un fenomeno sistemico che incide profondamente sulla distribuzione del potere, delle risorse, delle carriere e della legittimità scientifica.

Nel settore sanitario il *gender data bias* assume implicazioni par-

ticolarmente critiche, poiché le decisioni cliniche e organizzative sono sempre più supportate da sistemi di analisi automatizzata dei dati, algoritmi predittivi e strumenti di intelligenza artificiale. L'utilizzo di dati incompleti, non rappresentativi o costruiti su popolazioni prevalentemente maschili può produrre modelli predittivi distorti, con conseguenze dirette sull'equità dell'accesso alle cure, sull'accuratezza diagnostica e sulla qualità dei trattamenti erogati. Storicamente, molti studi clinici e database sanitari hanno sottorappresentato le donne o non hanno distinto adeguatamente tra sesso biologico e genere, contribuendo alla costruzione di standard diagnostici calibrati prevalentemente sul corpo maschile (Holdcroft, 2007; Criado Perez, 2019).

Le conseguenze di questo squilibrio possono essere significative. Ad esempio, sintomi, reazioni farmacologiche o fattori di rischio differenti tra uomini e donne rischiano di essere ignorati o sottovalutati, generando errori diagnostici, ritardi terapeutici o inapproprietezze cliniche. Come evidenzia Criado Perez, «when women are excluded from data, their needs become invisible» (2019, p. 29), con conseguenze significative in termini di maggiore esposizione a rischi sanitari. Anche questo fenomeno dimostra come il bias nei dati non sia soltanto una questione epistemologica, ma abbia effetti concreti sulla vita delle persone, contribuendo a generare disuguaglianze nell'accesso e nella qualità dell'assistenza sanitaria.

Un caso emblematico riguarda gli eventi avversi da farmaci: alcuni studi mostrano che le donne sperimentano più frequentemente effetti collaterali perché dosaggi e sperimentazioni sono spesso definiti sulla base di campioni maschili (Yu *et al.*, 2016; Zucker e Prendergast, 2020). Analogamente, nei sistemi di supporto decisionale automatizzato, dataset sbilanciati possono portare a sottostimare il rischio clinico di alcuni gruppi sociali oppure a sovrastimarli per altri⁶.

⁶ Alcuni algoritmi utilizzati negli ospedali statunitensi per prevedere quali pazienti necessitassero di cure aggiuntive o programmi di assistenza intensiva hanno mostrato una sistematica sottostima del rischio clinico per i pazienti afroamericani. Il sistema utilizzava come indicatore principale i costi sanitari sostenuti in passato: poiché, a causa delle disuguaglianze nell'accesso alle cure, i pazienti neri storicamente avevano sostenuto spese sanitarie inferiori rispetto ai pazienti bianchi, l'algoritmo interpretava erroneamente questo dato come indice di minore gravità clinica. Di conseguenza, molti pazienti afroamericani con condizioni di salute comparabili a quelle

Díaz-Rodríguez *et al.* (2023), analizzando i modelli epidemiologici sviluppati durante la pandemia da Covid-19, evidenziano come le differenze osservate tra uomini e donne non possano essere interpretate esclusivamente come effetti biologici, ma siano influenzate anche da fattori sociali, lavorativi, economici e culturali. Gli autori sottolineano che «failing to adapt the model to cater for those individual and cultural differences hinges both the accuracy of the predictions and put the group of individuals under a threat of disparate impact» (ivi, p. 2). In altri termini, ignorare le variabili sociali e strutturali nei modelli predittivi può produrre discriminazioni indirette, penalizzando specifici gruppi di popolazione nell'accesso alle risorse sanitarie o nella valutazione del rischio.

Come visto, la questione assume una rilevanza ancora maggiore se considerata in prospettiva intersezionale. Genere, etnia, età, condizione socioeconomica e provenienza geografica interagiscono infatti nel determinare esposizione ai rischi, possibilità di cura e qualità dell'assistenza ricevuta. Sistemi algoritmici addestrati su dati parziali rischiano dunque di automatizzare disuguaglianze già presenti nei sistemi sanitari, rafforzando vulnerabilità pregresse invece di ridurle. In questo senso, il bias nei dati non rappresenta soltanto un problema tecnico o epistemologico legato alla qualità della conoscenza prodotta, ma si traduce in effetti concreti sulle politiche pubbliche, sulla distribuzione delle risorse e sulla tutela del diritto alla salute.

La dimensione intersezionale del *gender data bias* rappresenta, come detto, un generale elemento di complessità. Le disuguaglianze non si distribuiscono in modo uniforme, ma si concentrano in corrispondenza dell'intersezione tra genere e altre categorie sociali, come la razza, la classe e l'età. Ciò implica che gli effetti del bias siano particolarmente intensi per i gruppi che occupano posizioni multiple di svantaggio. In questo senso, il *gender data bias* non può essere analizzato isolatamente, ma deve essere considerato come parte di un sistema più ampio di disuguaglianze interconnesse.

dei pazienti bianchi venivano classificati come meno bisognosi di assistenza specialistica. In questo caso, il dataset non rifletteva il reale stato di salute della popolazione, ma le disuguaglianze pregresse nell'accesso ai servizi sanitari, producendo così una distorsione discriminatoria nelle decisioni automatizzate (Obermeyer *et al.*, 2019).

Un altro campo di riflessione interessante è quello dell'istruzione. In questo caso, processi attraverso cui dati, indicatori statistici, sistemi di valutazione e tecnologie educative, pur presentandosi come neutrali e oggettivi, finiscono per incorporare presupposti culturali e sociali che riflettono prevalentemente esperienze maschili.

Nel campo educativo, il *gender data bias* emerge innanzitutto nei sistemi di valutazione standardizzati. Le prove scolastiche e i test comparativi internazionali tendono spesso a misurare competenze attraverso modelli culturali che riproducono aspettative stereotipate sui comportamenti di ragazze e ragazzi. Numerosi studi hanno mostrato, ad esempio, come le ragazze ottengano performance inferiori in matematica non necessariamente per una minore preparazione, ma per effetto dello *stereotype threat*, ossia della pressione psicologica derivante dalla consapevolezza di stereotipi negativi relativi alle capacità femminili nelle discipline scientifiche (Steele, 1997; Spencer *et al.*, 1999). In questo senso, il dato scolastico non rileva semplicemente una differenza “naturale”, ma contribuisce a produrla e consolidarla.

Un ulteriore ambito riguarda la produzione delle statistiche educative e degli indicatori di successo scolastico (Porter, 1995; Espeland e Stevens, 2008). Per lungo tempo, i dataset educativi hanno privilegiato variabili facilmente misurabili – voti, frequenza, abbandono scolastico – trascurando dimensioni più difficili da quantificare ma decisive per comprendere le disuguaglianze di genere, come il clima scolastico, le discriminazioni simboliche, le molestie o il carico di lavoro di cura che grava maggiormente sulle ragazze. Questa apparente neutralità statistica ha contribuito a rendere invisibili molte forme di esclusione. La questione assume oggi una rilevanza ancora maggiore con la diffusione delle piattaforme digitali e dei sistemi di intelligenza artificiale applicati all'istruzione. Gli algoritmi utilizzati per l'orientamento scolastico, l'apprendimento adattivo o la valutazione automatica apprendono infatti da dati storici che incorporano disuguaglianze pregresse. Se i dati di partenza riflettono una minore presenza femminile nelle discipline STEM, il rischio è che gli stessi sistemi digitali suggeriscano più frequentemente percorsi tecnico-scientifici ai ragazzi, rafforzando stereotipi già esistenti (Noble, 2018; UNESCO, 2020). In questo caso il bias non deriva soltanto da intenzioni discriminatorie esplicite, ma dall'uso di dati costruiti in contesti sociali segnati da asimmetrie di genere.

Anche i contenuti curriculari e i materiali didattici rappresentano un'importante fonte di *gender data bias*. Numerose ricerche hanno mostrato come i libri di testo continuino a sotto-rappresentare le donne nei ruoli scientifici, politici e culturali, contribuendo a naturalizzare una visione androcentrica della conoscenza (Blumberg, 2008). In termini sociologici, il curriculum scolastico agisce così come dispositivo di riproduzione culturale, legittimando implicitamente gerarchie di genere attraverso la selezione di ciò che merita di essere insegnato e ricordato (Bourdieu e Passeron, 1970).

Riconoscere i limiti dei dati non significa rifiutare la quantificazione o le tecnologie educative, ma adottare un approccio critico capace di interrogare i processi attraverso cui i dati vengono raccolti, classificati e interpretati. Solo in questo modo è possibile costruire sistemi educativi realmente inclusivi e sensibili alle differenze di genere.

Un'altra importante dimensione in esame riguarda il mercato del lavoro, in cui l'adozione di sistemi automatizzati di selezione e valutazione ha evidenziato come i dati storici possano incorporare e riprodurre discriminazioni di genere (Munoz *et al.*, 2024). Gli algoritmi utilizzati nei processi di reclutamento apprendono da dataset che riflettono pratiche passate, spesso caratterizzate da una sottorappresentazione delle donne in determinati ruoli. In questo contesto, il *gender data bias* contribuisce a perpetuare la segregazione occupazionale, indirizzando uomini e donne verso differenti percorsi professionali e consolidando differenze salariali. D'Ignazio e Klein osservano che i sistemi di dati e algoritmi «often reinforce existing inequalities rather than challenge them» (2020, p. 66), sottolineando il rischio che le tecnologie digitali rafforzino strutture sociali preesistenti anziché trasformarle.

Infine, è importante sottolineare come il *gender data bias* contribuisca a una forma di “naturalizzazione” delle disuguaglianze. Poiché i dati e gli algoritmi sono spesso percepiti come oggettivi e neutrali, le decisioni basate su di essi tendono a essere considerate legittime e incontestabili. Questo conferisce alle disuguaglianze una patina di oggettività che ne rafforza la stabilità e ne rende più difficile la messa in discussione. In questo senso, come detto, il *gender data bias* non solo produce disuguaglianze, ma contribuisce anche a legittimarle, rendendole parte integrante del funzionamento ordinario dei sistemi sociali e tecnologici.

Alla luce di queste considerazioni, appare evidente che affrontare il *gender data bias* richiede un approccio che riconosca la natura sociale e politica dei dati. Solo attraverso una riflessione critica sulle condizioni di produzione e utilizzo dell'informazione sarà possibile sviluppare sistemi più equi e inclusivi, capaci di ridurre anziché amplificare le disuguaglianze esistenti.

6. Conclusioni

Alla luce dell'analisi sviluppata, il *gender data bias* emerge non soltanto come una distorsione empirica o metodologica, ma come una questione epistemologica e politica di primaria rilevanza. Una prima implicazione riguarda il limite intrinseco delle soluzioni puramente tecniche proposte dalla *data science mainstream*. Interventi quali il bilanciamento dei dataset, la ricalibrazione degli algoritmi o l'introduzione di metriche di *fairness*⁷, pur rappresentando strumenti utili, presuppongono implicitamente che il bias sia localizzabile nei dati o nei modelli. Tuttavia, come evidenziato lungo tutto il lavoro, il bias è il risultato di processi sociali più ampi, che precedono e attraversano la produzione stessa dei dati. In questo senso, l'approccio tecnico rischia di operare una forma di "riduzionismo algoritmico", che affronta le manifestazioni superficiali del problema senza intervenire sulle condizioni strutturali che lo generano.

Un secondo elemento critico riguarda la definizione stessa delle categorie analitiche. Come mostrato dalla sociologia della conoscenza e dagli studi di Hacking (1982), le categorie statistiche non si limitano a descrivere la realtà, ma contribuiscono a costruirla. Nel caso del genere, questo implica che le modalità di classificazione e misurazione non sono neutrali, ma riflettono specifiche concezioni culturali e istituzionali. Il rischio, in questo contesto, è che i sistemi in-

⁷ Le metriche di *fairness* sono indicatori statistici e matematici utilizzati per valutare se un algoritmo o un sistema di intelligenza artificiale produca decisioni eque oppure discriminatorie nei confronti di determinati gruppi sociali (ad esempio per genere, etnia, età o condizione socioeconomica). In altri termini, servono a rilevare se un algoritmo tratta gruppi diversi in modo comparabile; se alcuni gruppi vengono sistematicamente penalizzati; se gli errori dell'algoritmo sono distribuiti in modo equo.

formativi consolidino categorie rigide e binarie, contribuendo alla marginalizzazione di soggettività non conformi e limitando la capacità di cogliere la complessità delle identità contemporanee.

La questione del potere assume qui un ruolo centrale. Seguendo la prospettiva foucaultiana, i dati possono essere interpretati come dispositivi di governo, strumenti attraverso cui si esercita una forma di potere che opera tramite la classificazione, la normalizzazione e la previsione. Gli algoritmi, in particolare, rappresentano una forma di potere “infrastrutturale”, che agisce in modo diffuso e spesso invisibile, orientando comportamenti e decisioni senza esplicitare i propri criteri. Questa opacità rende particolarmente difficile individuare le responsabilità e sviluppare forme efficaci di *accountability*, ponendo sfide rilevanti sul piano etico e regolativo.

Un ulteriore livello di complessità è introdotto dalla dimensione intersezionale del bias. Le evidenze empiriche mostrano chiaramente che le disuguaglianze non si distribuiscono in modo uniforme, ma si concentrano in corrispondenza dell’intersezione tra diverse categorie sociali. Questo implica che approcci focalizzati esclusivamente sul genere rischiano di produrre soluzioni parziali o addirittura controproducenti. La prospettiva intersezionale (Crenshaw, 1989) suggerisce invece la necessità di analizzare il bias come un fenomeno multidimensionale, che richiede strumenti teorici e metodologici capaci di cogliere la complessità delle relazioni sociali.

In questo quadro, assume particolare rilevanza il concetto di responsabilità epistemica (Code, 1987; Fricker, 2007). Se i dati sono socialmente costruiti, allora anche la loro interpretazione richiede una riflessione critica sui presupposti che li rendono possibili. Ciò implica riconoscere che la produzione della conoscenza non è un processo neutrale, ma un’attività situata, che comporta scelte, esclusioni e implicazioni normative. La sociologia, in questo senso, svolge un ruolo fondamentale nel rendere visibili tali dimensioni, contribuendo a sviluppare una maggiore consapevolezza critica.

Da queste considerazioni emerge la necessità di sviluppare una *data science* critica, capace di integrare competenze tecniche e riflessione sociologica. Tale approccio non si limita a correggere i bias esistenti, ma mira a ripensare l’intero processo di produzione dei dati, introducendo principi di inclusività, trasparenza e partecipazione. Questo implica, ad esempio, coinvolgere attivamente i gruppi sotto-

rappresentati nella definizione delle categorie e nella raccolta dei dati, nonché promuovere forme di *governance* che garantiscano maggiore *accountability* nei sistemi automatizzati.

Infine, il *gender data bias* solleva interrogativi rilevanti per le politiche pubbliche. In contesti come la sanità, il lavoro e l'istruzione, l'uso di dati distorti può contribuire a rafforzare disuguaglianze già esistenti, producendo effetti cumulativi nel tempo. Di conseguenza, le politiche basate sui dati devono essere accompagnate da una riflessione critica sulle condizioni di produzione delle informazioni utilizzate. Ciò richiede un dialogo interdisciplinare tra sociologia, statistica, informatica e scienze politiche, al fine di sviluppare strumenti analitici più sensibili alle differenze sociali.

In conclusione, il *gender data bias* non rappresenta un'anomalia del sistema, ma una sua componente strutturale, che riflette e riproduce le disuguaglianze sociali. Affrontarlo richiede un cambiamento di paradigma che vada oltre la dimensione tecnica, riconoscendo la natura profondamente sociale dei dati e promuovendo pratiche di produzione della conoscenza più eque e inclusive. Solo in questa prospettiva sarà possibile sviluppare sistemi informativi realmente capaci di rappresentare la complessità delle società contemporanee.

Riferimenti bibliografici

- Berger P.L., Luckmann T. (1966), *The Social Construction of Reality*, Anchor Books, New York.
- Blumberg R.L. (2008), *The Invisible Obstacle to Educational Equality: Gender Bias in Textbooks*, UNESCO, Paris.
- Bond R., Gibson C.M., Talasaz A.H., Van Tassell B.W. (2024), "Analysis of Gender and Race in Pharmacy Faculty and Administrators", *American Journal of Pharmaceutical Education*, 88, 1, <https://doi.org/10.1016/j.ajpe.2023.100585>
- Bourdieu P. (1979), *La distinction*, Minuit, Paris.
- Bourdieu P., Passeron J.C. (1970), *La reproduction*, Minuit, Paris.
- Bourdieu P., Chamboredon J.C., Passeron J.C. (1968), *Le métier de sociologue*, Mouton, Paris.
- Boyd D., Crawford K. (2012), "Critical Questions for Big Data", *Information, Communication & Society*, 15, 5, pp. 662–679, DOI: 10.1080/1369118X.2012.678878.

- Code L. (1987), *Epistemic Responsibility*, University Press of New England Hanover, NH.
- Crenshaw K. (1989), *Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics*, University of Chicago Legal Forum, 1989, 1, pp. 139–167, <https://chicagounbound.uchicago.edu/cgi/viewcontent.cgi?article=1052&context=uclf>.
- Crenshaw K. (1991), “Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color”, *Stanford Law Review*, 43, 6, pp.1241–1299, <https://doi.org/10.2307/1229039>
- Criado Perez C. (2019), *Invisible Women: Exposing Data Bias in a World Designed for Men*, Chatto & Windus, London.
- D’Ignazio C., Klein L.F. (2020), *Data Feminism*, MIT Press, Cambridge, MA.
- Desrosières A. (1993), *La politique des grands nombres. Histoire de la raison statistique*, La Découverte Paris.
- Díaz-Rodríguez N., Binkyte R., Bakkali W., Bookseller S., Tubaro P., Bacevičius A., Zhioua S., Chatila R. (2023), “Gender and sex bias in COVID-19 epidemiological data through the lens of causality”, *Information Processing and Management*, 60, 3, 103276, pp. 1–18, <https://doi.org/10.1016/j.ipm.2023.103276>
- Espeland W.N., Stevens M.L. (2008), “A Sociology of Quantification”, *European Journal of Sociology*, 49, 3, pp. 401–436, DOI:10.1017/S0003975609000150.
- Folbre N. (2006), “Measuring Care: Gender, Empowerment, and the Care Economy” *Journal of Human Development*, 7, 2, pp. 183–199, <https://doi.org/10.1080/14649880600768512>
- Foucault M. (1975), *Surveiller et punir. Naissance de la prison*, Gallimard, Paris.
- Fricker M. (2007), *Epistemic Injustice: Power and the Ethics of Knowing*, Oxford University Press, Oxford.
- Friedman B., Nissenbaum H. (1996), “Bias in Computer Systems”, *ACM Transactions on Information Systems*, 14, 3, pp. 330–347, <https://doi.org/10.1145/230538.230561>
- Gardner Z., Mooney P., De Sabbata S., Dowthwaite L. (2020), “Quantifying gendered participation in OpenStreetMap: Responding to Theories of Female (under) Representation in Crowdsourced Mapping”, *Geo-Journal*, 85, pp. 1603–1620, <https://www.jstor.org/stable/48730609>
- Gillespie T. (2014), *The Relevance of Algorithms*, pp. 167–194 in Gillespie T., Boczkowski P.J. and Foot K.A., eds., *Media Technologies: Essays on Communication, Materiality, and Society*, MIT Press Cambridge, MA.

- Gitelman L. (2013), *“Raw Data” Is an Oxymoron*, MIT Press, Cambridge, MA.
- Hacking I. (1982), “Biopower and the Avalanche of Printed Numbers”, *Humanities in Society*, 5, pp. 279–295.
- Harding S. (1986), *The Science Question in Feminism*, Cornell University Press, Ithaca.
- Holdcroft A. (2007), “Gender Bias in Research: How Does It Affect Evidence Based Medicine?”, *Journal of the Royal Society of Medicine*, 100, 1, pp. 2–3. <https://doi.org/10.1177/014107680710000102>
- Kahneman D. (2011), *Thinking, Fast and Slow*, Farrar, Straus and Giroux, New York.
- Kitchin R. (2014), *The Data Revolution*, Sage, London.
- Lam S.K., Uduwage A., Dong Z., Sen S., Musicant D.R., Terveen L., Riedl J. (2011), “WP: Clubhouse? An exploration of Wikipedia’s gender imbalance”, *Proceedings of the 7th International Symposium on Wikis and Open Collaboration* (WikiSym ‘11), pp. 1–10. <https://doi.org/10.1145/2038558.2038560>
- Latour B. (1987), *Science in Action*, Harvard University Press, Cambridge, MA.
- Latour B., Woolgar S. (1979), *Laboratory Life: The Construction of Scientific Facts*, Sage, Beverly Hills.
- MacNeill L., Driscoll A., Hunt A.N. (2015), “What’s in a Name: Exposing Gender Bias in Student Ratings of Teaching”, *Innovative Higher Education*, 40, 4, pp. 291–303, <https://doi.org/10.1007/s10755-014-9313-4>
- Munoz I., Kim P., O’Neil C., Dunn M., Sawyer S. (2024), “Platformization of Inequality: Gender and Race in Digital Labor Platforms”, *Proceedings of the ACM on Human-Computer Interaction*, 8, CSCW1, pp. 1–22. <https://doi.org/10.1145/3637385>
- Noble S.U. (2018), *Algorithms of Oppression: How Search Engines Reinforce Racism*, NYU Press, New York.
- O’Neil C. (2016), *Weapons of Math Destruction*, Crown, New York.
- Obermeyer Z., Powers B., Vogeli C., Mullainathan S. (2019), “Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations”, *Science*, 366, 6464, pp. 447–453, <https://doi.org/10.1126/science.aax2342>
- Pasquale F. (2015), *The Black Box Society: The Secret Algorithms That Control Money and Information*, Harvard University Press, Cambridge, MA.
- Pearl J. (2009), *Causality: Models, Reasoning, and Inference* (2nd ed.). Cambridge University Press, Cambridge, MA.
- Picardi I. (2019), “The Glass Door of Academia: Unveiling New Gendered Bias in Academic Recruitment”, *Social Sciences*, 8, 5, 160, <https://doi.org/10.3390/socsci8050160>

- Porter T.M. (1995), *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*, Princeton University Press Princeton.
- Potvin G., Hazari Z. (2016), “Student Evaluations of Physics Teachers: On the Stability and Persistence of Gender Bias”, *Physical Review Physics Education Research*, 12, 2, 020107, DOI: <https://doi.org/10.1103/PhysRevPhysEducRes.12.020107>, <https://journals.aps.org/prper/abstract/10.1103/PhysRevPhysEducRes.12.020107>
- Rodriguez-Seijas C., Morgan T.A., Zimmerman M. (2024), “Transgender and Gender Diverse Patients Are Diagnosed with Borderline Personality Disorder More Frequently Than Cisgender Patients Regardless of Personality Pathology”, *Transgender Health*, 9, 6, pp. 553–564. <https://doi.org/10.1089/trgh.2023.0062>
- Scheuerman M.K., Paul J.M., Brubaker J.R. (2019), “How Computers See Gender: An Evaluation of Gender Classification in Commercial Facial Analysis and Image Labeling Services”, *Proceedings of the ACM on Human-Computer Interaction*, 3, CSCW, 144, pp. 1–33. <https://doi.org/10.1145/3359246>
- Schwemmer C., Knight C., Bello-Pardo E.D., Oklobdzija S., Schoonvelde M., Lockhart J.W. (2020), “Diagnosing Gender Bias in Image Recognition Systems”, *Socius: Sociological Research for a Dynamic World*, 6, pp. 1–17, <https://doi.org/10.1177/2378023120967171>
- Spencer S.J., Steele C.M., Quinn D.M. (1999), “Stereotype Threat and Women’s Math Performance”, *Journal of Experimental Social Psychology*, 35, 1, pp. 4–28, <https://doi.org/10.1006/jesp.1998.1373>
- Steele C.M. (1997), “A Threat in the Air. How Stereotypes Shape Intellectual Identity and Performance”, *American Psychologist*, 52, 6, pp. 613–629, <https://doi.org/10.1037/0003-066X.52.6.613>
- Stephens M. (2013), “Gender and the GeoWeb: Divisions in the Production of User-Generated Cartographic Information”, *GeoJournal*, 78, 6, pp. 981–996. <https://doi.org/10.1007/s10708-013-9492-z>
- Tversky A., Kahneman D. (1974), “Judgment under Uncertainty: Heuristics and Biases”, *Science*, 185, 4157, pp. 1124–1131, <https://doi.org/10.1126/science.185.4157.1124>
- UNESCO (2020), *Artificial Intelligence and Gender Equality*, UNESCO, Paris.
- Valencia E. (2022), “Gender-Biased Evaluation or Actual Differences? Fairness in the Evaluation of Faculty Teaching”, *Higher Education*, 83, 6, pp. 1315–1333. <https://doi.org/10.1007/s10734-021-00744-1>
- Yoder-Himes D.R., Asif A., Kinney K., Brandt T.J., Cecil R.E., Himes P.R., Cashon C., Hopp R.M.P., Ross E. (2022), “Racial, Skin Tone, and

- Sex Disparities in Automated Proctoring Software”, *Frontiers in Education*, 7, 881449, pp. 1–16, <https://doi.org/10.3389/feduc.2022.881449>
- Yu Y., Chen J., Li D., Wang L., Wang W., Liu H. (2016), “Systematic Analysis of Adverse Event Reports for Sex Differences in Adverse Drug Events”. *Scientific Reports*, 6, 24955, <https://doi.org/10.1038/srep24955>
- Zheng X., Chen J., Yan E., Ni C. (2023), “Gender and Country Biases in Wikipedia Citations to Scholarly Publications”, *Journal of the Association for Information Science and Technology*, 74, 2, pp. 219–233, <https://doi.org/10.1002/asi.24723>
- Zuboff S. (2019), *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*, PublicAffairs, New York.
- Zucker I., Prendergast B.J. (2020), “Sex Differences in Pharmacokinetics Predict Adverse Drug Reactions in Women”, *Biology of Sex Differences*, 11, 32, pp. 1–14, <https://doi.org/10.1186/s13293-020-00308-5>