

T4F Series
PAPER N. 3
a.a. 2022/2023

**L'intelligenza artificiale in
campo biomedico:
medicina, AI e genere**

MARIA CHIARA PAGANI, ALTEA
ORICCHIO, FLAVIA LO IACONO

Trento BiLaw Selected Student Papers

I paper sono stati selezionati a conclusione del corso *Diritto e Intelligenza Artificiale* a.a. 2022-2023, organizzato all'interno della Cattedra Jean Monnet "T4F – TrAlning 4 Future. Artificial Intelligence and EU Law", coordinato presso l'Università di Trento dal docente Carlo Casonato.

L'intelligenza artificiale in campo biomedico: medicina, AI e genere

Maria Chiara Pagani, Altea Oricchio, Flavia Lo Iacono*

ABSTRACT: The paper aims to analyze critical issues of the use and misuse of gender-biased data sets for the training of biomedical artificial intelligence algorithms and to propose possible solutions. As far as the legal regulation of the artificial intelligence systems used in the testing, treatment and healthcare assistance phase is concerned, the authors highlight the crucial need of a flexible and risk-based regulation proposal, to seize the specific peculiarities of the algorithms involved. The regulation should provide for mandatory human oversight, the explainability of the output and the use of training “representative” data sets.

KEYWORDS: artificial intelligence; gender-specific medicine; gender; bias; AI Act.

SOMMARIO: 1. Introduzione – 2. La medicina di genere: una medicina non solo a misura d'uomo – 3. Algoritmi intelligenti e *bias* di genere – 4. I *gender bias* nell'analisi dei *Big Data* biomedici: l'esempio dei biomarcatori digitali – 5. “GP at hand” e le possibili soluzioni per neutralizzare i *bias* di genere – 6. La “*regulatory challenge*”: cenni all'AI Act e alla legge 11 gennaio 2018, n. 3 – 7. Conclusioni e raccomandazioni

1. Introduzione

Negli ultimi anni è cresciuta esponenzialmente l'attenzione circa l'impiego dei cd. “*Big data*”, per la ricerca e l'applicazione dell'intelligenza artificiale in campo biomedico. I *data set* utilizzati a questi fini includono variabili quali sesso e genere, che comportano i principali *bias* nel risultato (output).

Gli stereotipi di genere possono condurre a un'errata diagnosi, alla mancata previsione di effetti collaterali di un determinato farmaco su un soggetto appartenente ad un dato sesso/genere, a gravi conseguenze per la salute e il benessere dei pazienti¹. Il problema dei *bias* di sesso e genere riguarda sia i pazienti che non si riconoscono perfettamente nella “norma”², come ad esempio gli appartenenti alla comunità LGBTQUIA+ (soprattutto la comunità *intersex*, *gender non conforming* e *transgender*), sia quelle categorie di persone che non vengono coinvolte nelle fasi di studio e sperimentazione clinica, come donne, minoranze etniche, bambini in età pediatrica.

* Studentessa dell'Università degli Studi di Trento, Facoltà di Giurisprudenza.

¹ B. HEALY, *The Yentl Syndrome*, in *New England Journal of Medicine*, 325, 1991, 274-276

² K. ALBERT, M. DELANO, *Sex trouble: Sex/Gender slippage, sex confusion, and sex obsession in machine learning using electronic health record*, in *CellPress*, August 12, 2022.

Le autrici si propongono di analizzare le criticità generate dall'utilizzo di *data set gender-biased* per l'addestramento di algoritmi di intelligenza artificiale impiegati in campo biomedico e prospettare alcuni possibili rimedi.

Dopo una breve introduzione alla medicina di genere e alle tipologie di *bias* e dati inerenti all'intelligenza artificiale, si vedranno le caratteristiche dell'app di *digital health "GP at hand"* e le soluzioni ai *bias* prospettate dal team di Babylon, l'azienda britannica proprietaria e sviluppatrice dell'applicazione.

Per quanto concerne la regolazione giuridica dei sistemi di intelligenza artificiale impiegati nella fase di sperimentazione, cura e assistenza sanitaria, riteniamo di cruciale importanza la proposta di una regolazione flessibile, *risk-based* e in grado di cogliere le peculiarità degli algoritmi impiegati, che preveda come obbligatorio lo *human oversight*, l'*explainability* dell'output e l'impiego di *data set* di addestramento "rappresentativi". A questi fini sarà fatto qualche cenno all'Artificial Intelligence Act³ e alla legge n. 3 del 2018⁴.

Il presente lavoro individua quale fonte principale dei *bias* di genere in medicina l'utilizzo di *data set* di addestramento che già incorporano errori a causa della mancata inclusione nei campioni utilizzati per i test clinici e di laboratorio di categorie quali donne (specialmente donne in gravidanza e in età fertile), persone appartenenti alla comunità LGBTQUIA+, minoranze etniche, bambini in età pediatrica.

Seguiranno le proposte di soluzione che riteniamo possano neutralizzare o, al contrario, sfruttare i *gender bias* affinché la "*medical AI*" diventi un potente alleato della medicina genere-specifica.

Proponiamo non soltanto di coinvolgere già dalla fase di sperimentazione del farmaco o del trattamento gli appartenenti alle suddette categorie sottorappresentate, ma anche di assegnare loro un ruolo attivo nella progettazione degli algoritmi e di applicare alle soluzioni di intelligenza artificiale la teoria femminista del linguaggio, volta a neutralizzare gli stereotipi di genere codificati nei dati testuali.

2. La medicina di genere: una medicina non solo a misura d'uomo

È il 1991 quando Bernardine Healy, la prima donna ad assumere la carica di direttrice dei *National Institutes of Health* (NIH) statunitensi, denuncia la sottorappresentazione delle pazienti donne nella scienza medica, pubblicando l'articolo "*The Yentl Syndrome*"⁵.

³ Proposta di Regolamento del Parlamento europeo e del Consiglio n. 8115/2021 che stabilisce regole armonizzate sull'intelligenza artificiale (legge sull'intelligenza artificiale) e modifica alcuni atti legislativi dell'Unione (cd. AI Act).

⁴ Legge 11 Gennaio 2018, n. 3 Delega al Governo in materia di sperimentazione clinica di medicinali nonché disposizioni per il riordino delle professioni sanitarie e per la dirigenza sanitaria del Ministero della salute.

⁵ B. HEALY, *The Yentl Syndrome*, in *New England Journal of Medicine*, 325, 1991, 274-276, in <https://www.nejm.org/doi/full/10.1056/NEJM199107253250408>.

Rispetto agli uomini, le donne vengono ricoverate meno frequentemente e di rado sono sottoposte a specifiche indagini diagnostiche come coronarografie, *stent*, *bypass*, trombolisi. Tuttavia, il dato che solleva più criticità è che le donne sono quasi sempre escluse dalle sperimentazioni cliniche volte a introdurre nuovi farmaci, nuove tecnologie diagnostiche e terapeutiche⁶.

Questa “asimmetria della ricerca medica”⁷ ha condotto a gravi pregiudizi per la salute e il benessere di pazienti che hanno reagito in modo estremamente negativo alla somministrazione di taluni farmaci o trattamenti, elaborati e approvati sulla base di sperimentazioni cliniche avvenute esclusivamente su campioni di sesso maschile.

La talidomide, ad esempio, era un antidepressivo e sedativo commercializzato e diffuso tra il 1957 e il 1962 in diversi paesi, che ha condotto al più grande scandalo sanitario del mondo. Il farmaco, sviluppato dalla tedesca Chemie Grunethal GmbH, prima di essere immesso nel mercato era stato testato soltanto su animali di sesso maschile e uomini (spesso soldati). Nonostante l’evidente mancanza di sperimentazioni su pazienti di sesso femminile, i medici lo prescrissero a donne in gravidanza per ridurre gli attacchi d’ansia. L’esito fu tragico: la talidomide ha provocato la nascita di oltre 10.000 bambini affetti da gravissime malformazioni e la morte di 5.000 nati.

A seguito della provocazione lanciata dalla dottoressa Healy, nella comunità scientifica comincia farsi strada l’idea di una scienza in grado di cogliere e sfruttare le differenze esistenti tra uomini e donne: la medicina di genere o meglio: *gender-specific*. La medicina di genere consiste in un particolare approccio metodologico strettamente connesso alla pratica della “medicina di precisione” o “personalizzata”, di cui si riporta la definizione elaborata in seno al *National Research Council*: «*The tailoring of medical treatment to the individual characteristics of each patient (...) to classify individuals into subpopulations that differ in their susceptibility to a particular disease or their response to a specific treatment. Preventative or therapeutic interventions can then be concentrated on those who will benefit, sparing expense and side effects for those who will not*»⁸.

La professoressa Giovannella Baggio, titolare della cattedra di medicina di genere presso l’Università di Padova, afferma che “medicina di genere” significa «comprendere in che modo le malattie di tutti gli organi e sistemi si manifestino nei due generi e, soprattutto, valutare le differenze di genere rispetto ai sintomi delle malattie, alla necessità di differenti percorsi diagnostici e interpretazione dei risultati, alle differenze nella risposta ai farmaci o alla necessità di utilizzare farmaci diversi, alle differenze rispetto alla prevenzione di

⁶ E. GRIGLIÈ, G. ROMEO, *Per soli uomini: il maschilismo dei dati, dalla ricerca scientifica al design*, Torino, 2021.

⁷ *Ibidem*.

⁸ NATIONAL RESEARCH COUNCIL, *Toward precision medicine: building a knowledge network for biomedical research and a new taxonomy of disease*, in *National Academies Press*, Washington (DC), 2011, 12.

tutte le malattie»⁹. La medicina di genere non è una nuova specialità della scienza medica, ma consiste in un metodo basato sulla necessaria dimensione interdisciplinare della medicina, che è teso a studiare l'influenza delle variabili di sesso e genere sulla fisiologia, fisiopatologia e patologia umana. Inoltre, la medicina genere specifica si concentra sulle distinzioni sociali, culturali e territoriali per compiere un'analisi articolata e complessa che permette di raggiungere l'uguaglianza e l'equità in scelte mediche, politiche e sociali diversificate e personalizzate, che si adattano ai bisogni specifici delle persone e alla loro costituzione¹⁰.

La docente riporta che la prassi medica ormai codificata da linee guida è fondata su dati e prove scientifiche ottenuti su sperimentazioni avvenute solo su campioni di sesso maschile¹¹. Fra le cause di questa lacuna, rileva quella secondo cui i ricercatori sono sempre stati scettici nell'effettuare sperimentazioni cliniche su donne in età fertile, sia a tutela del loro benessere fisico sia per evitare di esporre le sperimentazioni alle variabili dipendenti dal ciclo mestruale. Un altro fattore alla base dell'esclusione delle donne dalle sperimentazioni cliniche si ravvisa nella variabilità biologica della donna o dell'animale di sesso femminile, che comporterebbe costi maggiori per i test clinici¹². Giustificazione, quest'ultima, che non persuade affatto, in quanto come non esiste la donna-tipo, non esiste nemmeno l'uomo – tipo.

L'EMA ha per decenni raccomandato, ma non imposto, di testare i medicinali anche su campioni di sesso femminile. L'AIFA si è espressa sul punto e ha stabilito che d'ora in avanti non approverà più protocolli se non vi saranno rappresentati i due generi/sessi¹³. Ciò non risolve il problema dei farmaci e dei trattamenti già approvati e già immessi nel mercato.

Le differenze che innegabilmente esistono a livello fisiologico tra i due sessi possono tradursi in differenze clinicamente rilevanti nella farmacocinetica e nella farmacodinamica¹⁴. Queste, insieme alla sottorappresentazione delle donne negli studi clinici, possono spiegare perché le donne riferiscano tipicamente più effetti collaterali rispetto agli uomini.

La farmacologa Flavia Franconi¹⁵ dell'Università di Sassari riporta l'esempio del farmaco anti-insonnia Ambien, in relazione al quale nel 2013 la FDA ha imposto di dimezzare il dosaggio del principio attivo. Tale decisione è stata adottata a seguito di indagini svolte per accertare la causa dell'incremento di incidenti

⁹ G. BAGGIO, *Dalla medicina di genere alla medicina gender-specific*, in *Ital J Gender-Specific Med*, 1, 2015, 3-5.

¹⁰ A. CARNEVALE, E. A. TANGARI, A. IANNONE, E. SARTINI, *Will Big Data and personalized medicine do the gender dimension justice?*, in *AI & Society*, 2, 2023, 829-841.

¹¹ G. BAGGIO, *op. cit.*, 3 ss.

¹² E. GRIGLIE', G. ROMEO, *op. cit.*

¹³ Terzo Congresso sulla Medicina di Genere, AIFA, Padova, 10-11 ottobre 2013.

¹⁴ D. CIRILLO, S. CATUARA SOLARZ, C. MOREY ET AL., *Sex and gender differences and biases in artificial intelligence for biomedicine and healthcare*, in *Digit. Med.*, 3, 2020, 81.

¹⁵ F. FRANCONI, I. CAMPESI, *Pharmacogenomics, pharmacokinetics and pharmacodynamics: interaction with biological differences between men and women*, in *British Journal of Pharmacology*, 171, 2014, 580-594.

stradali, che vedevano coinvolte donne che si addormentavano mentre erano alla guida perché non avevano smaltito la dose del farmaco assunto la sera precedente. Il motivo? Il dosaggio previsto per l'assunzione del farmaco era stato tarato esclusivamente sull'uomo. La stessa dinamica si ripropone con riguardo allo Zolpidem, un farmaco utilizzato nella cura dell'insonnia che, fra le altre cose, è noto anche come "droga dello stupro". Tuttavia, la questione non può circoscriversi al solo rapporto dosaggio-peso. La dott.ssa Franconi cita il caso dei tiazolidinedioni (TDZ), farmaci utilizzati nella cura dei pazienti diabetici per il controllo della glicemia, che purtroppo aumentano drasticamente il rischio di fratture ossee nelle donne di età superiore ai sessant'anni. Rispetto al caso di specie, l'AIFA ha sottolineato la necessità di studiare formulazioni su misura per pazienti di sesso femminile.

Un altro esempio ancora si ravvisa con riguardo alla statina, utilizzata nella cura del colesterolo e nella prevenzione primaria e secondaria per ridurre il rischio di infarto. La dottoressa Franconi afferma: «Quando le statine furono introdotte sul mercato negli anni Ottanta, nello studio clinico non c'era neanche una donna». La fisiologia femminile è molto diversa da quella maschile e la scienza medica deve necessariamente considerare questo aspetto¹⁶. Il corpo femminile, ad esempio, contiene meno acqua e più grassi (lipidi) rispetto a quello maschile e questo genera conseguenze importanti nella scelta del farmaco da prescrivere a una paziente: se si somministra ad una donna una molecola che non è liposolubile, questa tenderà ad accumularsi nell'organismo femminile e a produrre effetti tossici. Inoltre, studi recenti dimostrano che non soltanto gli organismi, ma le stesse cellule che compongono gli organismi maschili e gli organismi femminili sono diverse e danno risposte differenti allo stesso farmaco o alla stessa patologia¹⁷.

Il sesso e il genere del paziente hanno ripercussioni anche sul piano della progettazione di dispositivi medici, come ad esempio i cuori artificiali. Il primo cuore artificiale, progettato e sviluppato dall'azienda francese Carmat nel 2014, era stato infatti disegnato per essere compatibile con l'86 % dei pazienti uomini e solo con il 20% delle pazienti donna¹⁸.

Se si consultano i manuali di medicina, il paziente-target è un maschio bianco di 70 chili, che non solo non rappresenta affatto la popolazione femminile, ma nemmeno rispecchia totalmente quella maschile¹⁹. Quando le donne vengono menzionate nei testi di medicina, le si considera come una variazione dello standard di riferimento. Secondo uno studio²⁰ che ha monitorato i libri di testo nelle principali università di medicina in UE, USA e Canada, su 16.329 immagini analizzate i corpi maschili vengono usati tre volte più

¹⁶ Intervista a Flavia Franconi, in <https://www.fondazioneveronesi.it/magazine/articoli/Cardiologia/medicina-di-genere-i-cambiamenti-lenti-della-ricerca-sui-farmaci>.

¹⁷ C. AINSWORTH, *Sex and the Single Cells*, in *Nature*, 550, 2017, 6-8.

¹⁸ T. C. NGUYEN, *The World's First True Artificial Heart*, in *Smithsonian Magazine*, 25 dicembre 2013.

¹⁹ E. GRIGLIE', G. ROMEO, *op. cit.*

²⁰ E. GRIGLIE', G. ROMEO, *op. cit.*

spesso per illustrare “parti neutre” come gambe, addome o testa. La specificità della donna è relegata a quella che in gergo è diventata la “*bikini view*”: seno e apparato riproduttivo.

L'approccio metodologico proposto dalla medicina di genere mira a utilizzare le variabili di sesso e genere per ottimizzare il trattamento terapeutico e la diagnosi dei pazienti. Tuttavia, con il crescente impiego di algoritmi di intelligenza artificiale in ambito biomedico, la medicina deve affrontare un'ulteriore sfida: quella dei *bias* di genere che vengono “appresi” dall'algoritmo e che, di conseguenza, sono riprodotti nell'*output*. Rimediare a queste distorsioni di risultato diventa ancora più necessario dal momento che i *data set* utilizzati per addestrare l'algoritmo, che già contengono *gender bias*, appartengono alla categoria dei cd. *Big Data*²¹, che derivano da molteplici fonti.

Ai fini della nostra trattazione, la fonte principale di *big data* è rappresentata dalle cartelle cliniche elettroniche.

3. Algoritmi intelligenti e *bias* di genere

Quando si parla di “*bias*” nell'intelligenza artificiale si fa riferimento al fenomeno per cui durante il processo di apprendimento automatico l'algoritmo conduce a risultati discriminatori nei confronti di determinate categorie di soggetti²². Sviluppare strumenti imparziali è difficile, soprattutto se si riconosce che il loro funzionamento si basa su algoritmi che riflettono gli stereotipi e i preconcetti di chi li ha progettati. Per citare le parole dell'esperta Kate Crawford: “Come tutte le tecnologie prima di essa, l'intelligenza artificiale rifletterà i valori dei suoi creatori”²³.

L'eccessiva rappresentazione maschile nella progettazione di sistemi di intelligenza artificiale potrebbe aver condotto all'annullamento dei progressi fatti in tema di parità di genere.

I *bias* possono essere incorporati negli algoritmi di IA direttamente o indirettamente. Nel primo caso si tratta di preconcetti, elementi socioculturali presenti negli sviluppatori delle tecnologie di IA. Ad esempio, gli stereotipi di genere vivono, si modificano e perpetuano nella società che li accoglie e sono in prima battuta veicolati dal linguaggio. Poiché la maggior parte dei *data set* di addestramento dell'IA presenta una componente testuale, ciò comporta l'immissione dei *bias* di genere nella macchina che apprende il preconcetto codificato nel dato testuale e lo replica nell'*output*.

²¹ A. CARNEVALE, E. A. TANGARI, A. IANNONE, E. SARTINI, *op. cit.*, 829 ss.

²² P. TRAVERSO, *Breve introduzione tecnica all'Intelligenza Artificiale*, in *DPCE online*, 1, 2022, 155 ss.

²³ S. LEAVY, *Gender bias in artificial intelligence: the need for diversity and gender theory in machine learning*, in *Proceedings of the 1st International Workshop on Gender Equality in Software Engineering*, Association for Computing Machinery, New York, 2018, 14–16.

Sono invece *bias* indiretti quelli che insorgono a causa dei limiti dell'algoritmo, che nei sistemi di *machine learning* sono in primo luogo riconducibili al cd. fenomeno della *black box*, che rende opaco il percorso seguito della macchina per giungere a un determinato risultato²⁴.

Con riguardo all'AI impiegata nella medicina di genere, un'altra distinzione rilevante all'interno dei *bias* sessuali e di genere è quella tra pregiudizi desiderabili e pregiudizi indesiderabili. Un pregiudizio desiderabile implica la considerazione e l'utilizzo consapevole delle differenze di sesso e di genere per fare una diagnosi precisa e mirata e raccomandare una terapia su misura e di conseguenza più efficace per ogni individuo. Un pregiudizio indesiderabile, invece, è quello che conduce alla discriminazione non intenzionale e non necessaria di una data categoria di soggetti sulla base di sesso e genere.

Le principali fonti di *bias* indesiderabili in materia di AI e medicina si ravvisano nella falsa rappresentazione target, che esclude le minoranze (es. persone appartenenti a minoranze etniche o razziali, persone *transgender*, persone *gender non conforming*, persone appartenenti alla comunità LGBTQUIA+) e nella mancanza di campioni rappresentativi della componente femminile della popolazione, soprattutto donne in età fertile e in gravidanza, ma anche bambini in età pediatrica, nei campioni utilizzati per la ricerca clinica psichiatrica.

Se i pregiudizi indesiderabili vengono introdotti nelle tecnologie di intelligenza artificiale perché sono presenti nei *data set* di addestramento, la macchina apprenderà il pregiudizio e, di conseguenza, restituirà un risultato discriminatorio e distorto. Questo fenomeno può essere in parte corretto nei sistemi di intelligenza artificiale *model-based*, lavorando sulla corretta selezione del *data set* con cui addestrare l'algoritmo e soprattutto avvalendosi dell'expertise del medico che, ad esempio, possiede gli strumenti e le conoscenze per riconoscere e correggere una diagnosi errata. Al contrario, in caso di *machine learning*, diventa ancora più difficile prevenire i *bias* indesiderabili perché è lo stesso algoritmo che, inspiegabilmente, li produce attraverso una discriminazione intrinseca alla macchina o sulla base di un principio irragionevole di selettività dei dati e delle informazioni.

4. I *gender bias* nell'analisi dei *Big Data* biomedici: l'esempio dei biomarcatori digitali

Negli ultimi anni, anche a causa della pandemia Covid-19, l'utilizzo di sistemi di IA nel settore biomedico è cresciuto esponenzialmente, soprattutto in virtù della spiccata versatilità di queste tecnologie e della loro capacità di processare una particolare categoria di dati e di elaborarli in tempo reale: i c.d. *Big Data*. I *big data* si caratterizzano per il loro volume, la velocità, la varietà, la veridicità e il valore: le cd. 5V dei *big data*²⁵.

²⁴ L. SCAFFARDI, *La medicina alla prova dell'intelligenza artificiale*, in *DPCE Online*, 1, 2022, 349-359.

²⁵ A. CARNEVALE, E. A. TANGARI, A. IANNONE, E. SARTINI, *op. cit.*, 829 ss.

I *big data* biomedici si ricavano principalmente dai sistemi di cartelle cliniche elettroniche, ma anche avvalendosi dei biomarcatori digitali, cioè indicatori fisiologici, psicologici e comportamentali basati su dati che includono interazioni uomo-computer (ad es. colpetti, digitazione, sfioramenti), attività fisica (passo, andatura, destrezza, equilibrio) e variazioni della voce, raccolti da strumenti indossabili (cd. *wearable devices*), portatili, impiantabili e in certi casi anche ingeribili. L'utilizzo di questi dispositivi può facilitare la diagnosi di una patologia, la valutazione degli effetti di un particolare trattamento sanitario, la prognosi di un paziente.²⁶ Inoltre, le informazioni ottenute tramite questi sensori possono essere riportate su dispositivi mobili come gli smartphones, che monitorano le condizioni dell'individuo a cui sono applicati, segnalando il caso in cui un parametro dovesse uscire dal "range di normalità".

Le alte potenzialità di questi dispositivi permettono il monitoraggio costante degli indici di salute di un individuo in modo più preciso e accurato rispetto ai questionari, ai *self-report* e ai test psicometrici che oggi si utilizzano nella pratica clinica. Purtroppo, anche i biomarcatori digitali si scontrano con l'esistenza di *bias* indesiderabili presenti nei *data set* di addestramento dei modelli che forniscono gli indicatori di salute. Se l'algoritmo di un biomarcatore digitale viene addestrato con *data set* in cui sono presenti per la maggior parte campioni di sesso maschile, questo strumento rileverà in maniera decisamente più accurata i sintomi che si manifestano in maniera più frequente nei pazienti di sesso maschile. In altri casi, come detto in precedenza, il *bias* indesiderato deriva dallo stesso dispositivo digitale.

Tra le varie fonti di pregiudizi indesiderati presenti nei biomarcatori, vi sono la disparità nell'accesso e utilizzo di dispositivi digitali da parte di uomini e donne, nonché i livelli di istruzione, il reddito, l'età²⁷.

L'analisi dei *big data* può essere un prezioso supporto per la medicina personalizzata, che si caratterizza per essere "precisa", "predittiva" e "partecipativa".²⁸ Tuttavia, poiché in medicina i *big data* vengono utilizzati a supporto di decisioni umane (quelle dei medici), occorre tener presente che questi, per quanto statisticamente veloci e voluminosi, non possono mai fare previsioni sempre realistiche poiché una previsione ricavata esclusivamente dai dati è un'operazione matematica che non tiene conto della dimensione umana propria delle previsioni: mancano le valutazioni personali e l'esperienza.

²⁶ *Ibid.*

²⁷ *Ibid.*

²⁸ L. HOOD, S.H. FRIEND, *Predictive, personalized, preventive, participatory (P4) cancer medicine*, in *Nat. Rev. Clin. Oncol.*, 2011, 184–187.

5. “GP at hand” e le possibili soluzioni per neutralizzare i *bias* di genere

La nostra riflessione parte dai rischi e dalle conseguenze generati da algoritmi non correttamente addestrati, con riguardo alle app di *digital health* pensate per tutti e tutte, ma progettate su una certa idea di uomo-tipo (maschio bianco di circa 70 chili). Eugenio Santoro (responsabile del laboratorio di informatica medica, Dipartimento di Salute Pubblica, Istituto di Ricerche Farmacologiche “Mario Negri” IRCCS) evidenzia la sottorappresentanza femminile in molti ambiti della *digital health*²⁹.

Nei paesi ad alto reddito come gli USA, i dati indicano che le donne sono più propense a utilizzare gli strumenti di *digital health* rispetto agli uomini. Tuttavia, la maggior parte degli strumenti disponibili non si rivolge a un pubblico femminile e gli uomini sono spesso considerati il target predefinito nel disegno dell’interfaccia e nella realizzazione delle funzioni. L’autore incalza sostenendo che il genere condiziona anche il disegno, il layout e il linguaggio di molte app per la salute. Uno studio condotto da “*Lancet Digital Health*”, rileva che mentre la maggior parte delle app per la salute rivolte agli uomini presentavano un design e un’interfaccia neutrale, le app per le donne, al contrario, erano caratterizzate da una scelta di design che rifletteva lo stereotipo femminile (colore rosa, cuori, fiori etc.). Un altro esempio può ravvisarsi nelle app di fitness rivolte agli uomini, che mirano a risultati legati alla forza e muscolatura, mentre le app rivolte alle donne si concentrano su magrezza e la sinuosità delle forme femminili.

L’app “GP at Hand” rilasciata nel Regno Unito da Babylon nel 2017 e utilizzata come alternativa al *National Health System* (NHS) britannico per suggerire ai pazienti diagnosi basate sui sintomi da loro riferiti, filtrare e distinguere casi e patologie, prenotare prestazioni o visite mediche (sia a distanza in modalità telematica sia tramite appuntamenti in presenza), è stata accusata funzionare sulla base di un algoritmo non correttamente addestrato³⁰. Tale accusa muove dal caso di una donna fumatrice di 59 anni che, dopo aver riferito i tipici sintomi di un attacco di cuore (presenza di dolore toracico, mancanza di respiro e ansia), ha ricevuto una diagnosi di attacco di panico; al contrario, ad un utente maschio che ha riportato gli stessi identici sintomi è stato diagnosticato l’attacco cardiaco. La *ratio* si scorge nei criteri di probabilità utilizzati dall’algoritmo dell’applicazione: dato che statisticamente le donne sembrano soffrire maggiormente di ansia e depressione rispetto agli uomini, all’app è sembrato più probabile che la donna fumatrice fosse in preda ad un attacco di panico e non ad un attacco di cuore.

Sorge spontanea una domanda: come eliminare o limitare i *bias* di genere nei dispositivi di *digital health*? La maggior parte dei ricercatori ritiene che il primo passo sia coinvolgere le donne nel processo di disegno,

²⁹ E. SANTORO, *Le (troppe) disparità di genere nella salute digitale*, in *HealthTech360*, 20 ottobre 2021, disponibile su: <https://www.healthtech360.it/salute-digitale/le-troppe-disparita-di-genere-nella-salute-digitale/>.

³⁰ BABYLON, *Ladies (and everyone else), let’s kick some bias*, in <https://www.babylonhealth.com/en-gb/blog/tech/ladies-and-gentlemen-lets-kick-some-bias>.

sviluppo e test delle app e degli strumenti di *digital health*. Ciò al fine di capire il contesto in cui vengono utilizzate, per che scopo e per renderle più accurate. Altri autori propongono di utilizzare *data set* di addestramento più imparziali. Tale soluzione non persuade nell'ottica dell'impiego di tecnologie di IA nel contesto della medicina di genere, dove le variabili di sesso e genere, costituendo *bias* desiderabili possono contribuire all'elaborazione di diagnosi e terapie sempre più adeguate alle specifiche caratteristiche ed esigenze dell'individuo-paziente.

A fronte delle accuse ricevute, Babylon ha affermato che il suo *Symptom Checker* non è uno strumento diagnostico e che non è progettato per prospettare lo scenario peggiore o più improbabile. Al contrario, si tratta di uno strumento sviluppato per fornire informazioni mediche sulle condizioni di salute che, statisticamente, corrispondono ai sintomi inseriti.

Occorre sottolineare che l'app lavora su tre livelli (1. *Symptom Checker*; 2. Visita telematica con un medico del team Babylon; 3. Visita in presenza con un medico del team Babylon) e che i dati utilizzati al primo livello per elaborare la diagnosi non sono soltanto i sintomi autoriferiti, ma anche i dati e gli indici di salute contenuti nella cartella clinica elettronica dell'utente.

La dottoressa Claudia Pastides, una fra le donne che rappresentano i 2/3 dei medici del team Babylon, spiega che il *Symptom Checker* opera sulla base di indici di probabilità. Esso effettua calcoli basati su differenti dati epidemiologici provenienti da un numero elevato di studi di ricerca, presentando quindi centinaia di possibili condizioni di corrispondenza, da cui visualizza quelle che corrispondono maggiormente ai sintomi inseriti.

La dottoressa prosegue sostenendo che a differenza di un medico umano, l'intelligenza artificiale non è influenzata dai casi che gli sono stati sottoposti durante le varie sessioni settimanali; perciò, è meno probabile che costruisca una diagnosi su casi che le sono stati sottoposti in precedenza. Inoltre, la dottoressa afferma che l'IA non ha opinioni su sesso e genere e che non guarda i dati che riceve attraverso il velo del pregiudizio. Il team di Babylon perviene alla conclusione secondo cui il medico dell'IA è l'unico a non essere offuscato dalla forza involontaria del pregiudizio di genere. A fronte di nuove critiche e con il concreto rischio di contraddirsi, il team di Babylon ha pubblicato un articolo³¹ controverso nel quale spiega la difficoltà per gli algoritmi di intelligenza artificiale di essere privi di *bias* di genere, che derivano dallo scarso coinvolgimento delle donne nei procedimenti di sviluppo di *machine learning*. Gli autori concludono, in via totalmente diplomatica, che i *data set* con cui addestrare gli algoritmi rappresentare campioni di persone di entrambi i sessi e di qualsiasi genere, etnia, età.

³¹ H. LEE, *Directions for sex and gender-based health research in Korea: implications of the Amendments of the Framework Act on Science and Technology*, in *Korean J. Women Health Nurs.*, 4, 2022, 269-274.

Tuttavia, vi sono casi in cui l'impiego dell'IA in medicina non soltanto conduce a risultati diagnostici e terapeutici più accurati, ma salva la vita. Un recente articolo pubblicato sul New York Times sottolinea l'importanza dell'utilizzo dell'intelligenza artificiale nell'individuare e riconoscere il cancro nelle sue varie forme. In Ungheria, dove si assiste al massiccio e fertile ricorso a sistemi di intelligenza artificiale per l'assistenza sanitaria, per la prima volta l'algoritmo ha superato l'esperienza dei medici. Ciò è accaduto con una donna che si è sottoposta a una mammografia, nella quale due radiologi non avevano individuato nulla di sospetto, mentre l'intelligenza artificiale aveva evidenziato con dei cerchi rossi delle potenziali cellule tumorali. Questo sembra essere un eccezionale passo avanti per la medicina, ma soprattutto per l'individuazione fin dagli stadi iniziali del tumore al seno. Il National Cancer Institute ha stimato che più del 20% dei tumori al seno non vengono individuati negli *screenings* della mammografia.

6. La “regulatory challenge”: cenni all’AI Act e alla legge 11 gennaio 2018, n. 3

Dal punto di vista della regolazione giuridica dell'utilizzo di dispositivi di intelligenza artificiale in ambito medico, si possono considerare due diversi strumenti normativi: la legge n. 3 del 2018 e l'AI Act.

La legge n. 3 del 2018³² è tesa a riformare le disposizioni in materia di sperimentazione clinica dei medicinali per uso umano e si coordina con il reg. UE n. 536/2014³³.

Ai fini della nostra trattazione, rilevano gli articoli 1 e 3 recanti le norme relative alla delega al Governo per il riassetto e la riforma della normativa in materia di sperimentazione clinica e l'applicazione e diffusione della medicina di genere nel Servizio sanitario nazionale.

Il comma 1 dell'art. 1 dispone che il governo adottare i decreti legislativi necessari per il «riassetto e la riforma delle disposizioni in materia di sperimentazione clinica dei medicinali» che contengano esplicito riferimento alla medicina di genere e all'età pediatrica. Al comma 2 lett. c, inoltre, si prevede la necessità dell'individuazione delle modalità per la migioria di centri clinici dedicati agli studi clinici, da «condurre con un approccio metodologico di medicina di genere»³⁴.

L'art. 3 al comma 1 prevede più in generale che, il Ministero della salute, sentita la Conferenza permanente per i rapporti tra lo Stato, le Regioni e le Province autonome di Trento e di Bolzano predisponga un piano volto «alla diffusione della medicina di genere mediante divulgazione, formazione e indicazione di pratiche sanitarie che nella ricerca, nella prevenzione, nella diagnosi e nella cura tengano conto delle differenze derivanti dal genere» con il fine di migliorare la precisione delle prestazioni erogate dal Servizio sanitario

³² Legge 11 Gennaio 2018, n. 3 Delega al Governo in materia di sperimentazione clinica di medicinali nonché disposizioni per il riordino delle professioni sanitarie e per la dirigenza sanitaria del Ministero della salute.

³³ Regolamento (UE) n. 536/2014 del Parlamento europeo e del Consiglio del 16 aprile 2014 sulla sperimentazione clinica di medicinali per uso umano.

³⁴ Art. 1, L. 11 Gennaio 2018, n. 3.

nazionale. A sostegno di quanto appena detto, alla lettera c dello stesso comma, si prevede la «promozione e sostegno dell'insegnamento della medicina di genere, garantendo adeguati livelli di formazione e di aggiornamento del personale medico e sanitario»³⁵.

Al comma 3 dell'art. 3 si stabilisce, inoltre, che il Ministro della salute ha il dovere di emanare apposite raccomandazioni, agli Ordini e ai Collegi delle professioni sanitarie alle società scientifiche e alle associazioni di operatori sanitari non iscritti a Ordini o Collegi, volte a promuovere l'applicazione della medicina di genere sul territorio nazionale. Mentre al comma 4 si predispone che lo stesso Ministero, in concerto con quello dell'istruzione, dell'università e della ricerca, proponga un Piano formativo nazionale per la medicina di genere «volto a garantire la conoscenza e l'applicazione dell'orientamento alle differenze di genere nella ricerca, nella prevenzione, nella diagnosi e nella cura»³⁶.

Infine, al comma 5 dell'art. 3, si dispone che il Ministero della salute trasmetta ogni anno alle Camere una relazione sulle azioni di promozione e di sostegno della medicina di genere attuate nel territorio nazionale, «anche attraverso l'istituzione di un Osservatorio dedicato alla medicina di genere istituito presso gli enti vigilati dal Ministero della salute»³⁷. Tuttavia, come emerge dalla *littera legis*, si tratterebbe di mere raccomandazioni e non di obblighi vincolanti (ad eccezione dell'adozione dei decreti legislativi di attuazione del reg. UE n. 536/2014).

L'AI Act, la proposta di regolamento dei sistemi di intelligenza artificiale promossa dalla Commissione Europea³⁸, mira a regolare l'IA attraverso il c.d. *risk-based approach*, che distingue diversi obblighi di conformità in base al grado di rischio che i software di intelligenza artificiale possono causare a danno dei diritti fondamentali degli individui. In ambito biomedico, se l'intelligenza artificiale non è ben regolata, il rischio è quello di ledere il diritto alla salute e il diritto alla privacy dei pazienti. Più alto è il rischio e più stringenti sono gli obblighi di compliance imposti agli sviluppatori, distributori e utilizzatori dei sistemi di intelligenza artificiale.

In campo biomedico, è cruciale quanto disposto dall'AI Act in tema di *Human Oversight* (art. 14) e *Data Governance* (art. 10)³⁹.

Innanzitutto, per evitare che gli algoritmi riproducano i *bias* di genere, i dati immessi come input devono essere rappresentativi. Questo comporta che i *data set* impiegati per l'addestramento dell'algoritmo devono

³⁵ Art. 3, L. 11 Gennaio 2018, n. 3.

³⁶ Art. 3, L. 11 Gennaio 2018, n. 3.

³⁷ Art. 3, L. 11 Gennaio 2018, n. 3.

³⁸ Proposta di regolamento del Parlamento europeo e del Consiglio n. 8115/2021 che stabilisce regole armonizzate sull'intelligenza artificiale (legge sull'intelligenza artificiale) e modifica alcuni atti legislativi dell'Unione.

³⁹ Art. 10 e art. 14 della Proposta di regolamento del Parlamento europeo e del Consiglio n. 8115/2021 che stabilisce regole armonizzate sull'intelligenza artificiale (legge sull'intelligenza artificiale) e modifica alcuni atti legislativi dell'Unione

necessariamente contenere campioni che siano rappresentativi di tutta la popolazione: donne, donne in età fertile e in gravidanza, bambini di età pediatrica, persone appartenenti a minoranze etniche, persone *gender non-conforming*, persone *transgender*, persone intersessuali e in generale coloro che si identificano nella comunità LGBTQUIA+. Più i data set sono diversificati e contengono le variabili dello spettro sessuale e di genere, minore sarà il rischio di errore e distorsione nell' output.

Con riguardo all'art. 14, riteniamo che la supervisione e il controllo umano sui sistemi di intelligenza artificiale impiegati in medicina sia fondamentale. Infatti, l'intelligenza artificiale in medicina può essere considerata un'arma a doppio taglio: da un lato può migliorare le prestazioni e l'assistenza sanitaria, dall'altro rischia di riprodurre *bias* indesiderati e generare diagnosi errate. È proprio nel momento in cui la tecnologia mostra i suoi limiti che è necessario l'intervento umano di chi possiede le conoscenze e gli strumenti propri della scienza biomedica: il medico.

La regolamentazione della IA medica potrebbe adottare un *neutral approach* per garantire l'equità nella generazione e nel trattamento dei dati, rimuovendo informazioni sensibili come il sesso e il genere e tutte le altre variabili connesse. Tuttavia, riteniamo che un approccio genere-specifico sia da privilegiare. Se esistono differenze intrinseche alla popolazione, quali sesso e genere, un approccio neutrale è fallace perché porterebbe a risultati imprecisi e non equi nei confronti di specifiche minoranze. Anziché promuovere la neutralità della medicina e della ricerca e più in generale la neutralità dei dati, è da preferirsi un approccio differenziato. Riteniamo che lo stesso approccio *diversity-based* debba essere adottato nella regolamentazione giuridica dell'intelligenza artificiale. In particolare, ricollegandoci a quanto stabilito dall'art. 14 dell'AI Act, occorre ripensare l'architettura dei big data generati dai modelli di *machine learning*: qui il software non decide autonomamente, ma consente a un essere umano dotato di una specifica esperienza in un dato campo, come ad esempio il medico, di decidere in base alle prove addotte dall'algoritmo e di integrare queste informazioni con il suo expertise. L'integrazione dei processi decisionali tecnologici con la supervisione e l'intervento di un essere umano qualificato può valorizzare la dimensione del sesso e del genere nelle tecnologie informatiche per l'assistenza sanitaria.

7. Conclusioni e raccomandazioni

Riconoscere i pregiudizi di genere nei dati di addestramento per gli algoritmi di intelligenza artificiale è senza dubbio un compito complesso e delicato, ma non impossibile.

Innanzitutto, occorre segnalare la crescente attenzione per la *gender equality* che sta investendo il mondo dell'IA. Promuovere la carriera delle donne e delle minoranze nel campo dell'Intelligenza Artificiale è essenziale per impedire che i progressi nell'uguaglianza di genere sostenuti da decenni di pensiero femminista vengano annullati.

Dall'analisi svolta, emerge una discriminazione di genere latente, che opera prima ancora che la ricerca medica produca i suoi risultati. Tali discriminazioni spesso agiscono su schemi inconsci, di precomprensione, che non sempre consentono ai ricercatori di interrogarsi sul rapporto tra scienza e società. Ciò espone al rischio di essere ciechi di fronte a quelle forme di ingiustizia che rendono vulnerabili alcuni gruppi sociali, gli stessi a cui appartengono i pazienti esclusi dalle sperimentazioni cliniche⁴⁰.

Tuttavia, anche quando la dimensione di genere viene presa in considerazione, viene trattata come un contrasto naturale e binario tra maschile e femminile. Come ha sottolineato Anne Fausto-Sterling⁴¹, le categorie binarie di sesso e genere sono una "decisione sociale" e non la rappresentazione di realtà biologiche inevitabili.

In conclusione, riteniamo che sia necessario distinguere tra *bias* indesiderabili e *bias* desiderabili per garantire l'inclusione dei *bias* desiderabili nello sviluppo delle tecnologie di IA.

Si raccomanda la partecipazione delle categorie sottorappresentate (donne, minoranze etniche, appartenenti alla comunità LGBTQUIA+, bambini in età pediatrica) sia ai trial clinici sia ai processi di sviluppo e progettazione dei sistemi di intelligenza artificiale per creare un paradigma medico-tecnologico *gender-oriented* e a misura di paziente. Da ultimo, riteniamo sia importante incorporare le teorie femministe del linguaggio volte ad eliminare gli stereotipi di genere presenti nel linguaggio umano che possono influenzare l'algoritmo attraverso l'introduzione di dati testuali distorti.

Occorre adoperarsi per una ricerca medica che fornisca una fotografia il più possibile rappresentativa della popolazione. I sistemi di intelligenza artificiale dovrebbero essere progettati per confrontare specifiche caratteristiche da un punto di vista intersezionale⁴², come il genere nella sua configurazione non binaria. Inoltre, gli sviluppatori dovrebbero utilizzare solo informazioni sensibili relative a genere, sesso o razza in applicazioni specifiche e regolamentate in cui è dimostrato che sono importanti, specialmente in campo medico. Per quanto concerne altre applicazioni, è da preferirsi un approccio neutrale.

⁴⁰ M. POT, W. SPAHL, B. PRAINSACK, *The gender of biomedical data: challenges for personalised and precision medicine*, in *Somatechnics*, 9, 2019, 170–187.

⁴¹ A. FAUSTO – STERLING, *Sexing the body: gender politics and the construction of sexuality*, New York, 2000.

⁴² K. CRENSHAW, *Demarginalizing the intersection of race and sex: a black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics.*, in *Univ. of Chicago Legal Forum*, 140, 1989, 139–167.