

bIAs: su computazione algoritmica e razzismo sistemico*

Emilia Marra

Chi raccoglie i frutti dell'innovazione tecnica? Quali corpi ne pagano il prezzo? A François Magendie si deve il motto del metodo sperimentale in ambito medico: «sperimentate, sperimentate, e il resto verrà da sé». Georges Canguilhem, epistemologo francese avvezzo a frugare nella storia delle discipline scientifiche, ricorda che, «accusato di praticare esperimenti sull'uomo, Magendie negò»¹: il suo laboratorio *en plein air* fu la vivisezione animale. Al tempo non mancarono certo proteste, animate forse però più da un timore di stampo egoistico che da una preoccupazione animalista *ante litteram*. Più che la compassione, la rivendicazione antispecista nascondeva una certa conoscenza della logica inferenziale: «giungere a delle conclusioni sull'uomo a partire dall'animale significa abolire la distanza tra l'uno e l'altro, significa una assimilazione ritenuta materialistica e in grado, a lungo andare, di suscitare la tentazione di una estensione all'uomo»². Perdi più, quel pericolo era, in effetti, non solo fondato, ma già in atto. Se sperimentare sugli esseri umani significa infatti somministrare loro farmaci in via di sperimentazione, allora non si può non convenire che Magendie abbia sperimentato nel suo laboratorio *indoor*, l'ospedale pubblico, che senza particolari restrizioni gli fornì, nel corso della prima metà dell'Ottocento, una popolazione che «permette la formazione di gruppi e il loro confronto»³. La cosiddetta razionalità scientifica si è nutrita dei corpi di coloro i quali il lessicografo francese Antoine Furetière definiva già nel 1690, nel suo *Dizionario universale*, «persone di poca importanza»⁴. A questa categoria appartiene ogni fuori norma, e il fuori norma si dice in molti modi: disabile, condannato a morte, orfano, prostituta, l'omosessuale, lo schiavo, il folle o l'indigeno, l'indigente e il nero. Grégoire Chamayou, filosofo francese oggi *chargé de recherche* presso l'ENS di Lione, parla a questo proposito molto opportunamente di «corpi vili» (giocando con la locuzione latina, spesso ironica, *in corpore vili*), materiale sperimentale per il quale

* Questo studio è stato condotto nel corso dell'assegno di ricerca "Le discriminazioni prodotte dalla profilazione algoritmica", di durata 12 mesi, presso il Dipartimento di Studi Umanistici dell'Università del Salento nell'ambito del progetto G.E.A. Global, Green, Generative and Equal Educational Activities (Codice AID 012618/03/8).

¹ G. Canguilhem, *L'effetto della batteriologia nella fine delle "teorie mediche"*, in Id., *Ideologia e razionalità nella storia delle scienze della vita*, tr. it. di E. Marra, Mimesis, Milano 2025, p. 90.

² Ibidem.

³ Id., p. 91.

⁴ Cfr. A. Furetière, *Dictionnaire universel*, Arnout et Reinier Leers, La Haye 1690, in G. Chamayou, *Les corps vils*, La Découverte, Paris 2014, p. 7.

la scienza moderna ha sempre mostrato una predilezione. I corpi vili sono corpi sacrificabili, luogo di esercizio privilegiato della violenza travestita da progresso tecnico, non proprio animali, ma nemmeno umani in senso proprio. I corpi vili sono accomunati dall'appartenenza a una sorta di commistione viva tra umano e non umano, tra soggetto di diritto e oggetto del potere. All'interno di questo territorio ibrido, spesso caratterizzato dal vuoto normativo, si apre la caccia al meticcio epistemico. Chi è oggi il meticcio, il liminale, a quale corpo si dà la caccia? Si tratta ancora del corpo dal quale, si potrebbe dire giocando con il celebre corso di Foucault (1976), *Il faut/fallait défendre la société*, bisogna difendere la società? E una volta individuato il meticcio, il dispensabile, il corpo di poco conto, quali sono le regole di diritto messe in opera per produrre discorso scientifico? Una intuizione che mi sembra particolarmente fruttuosa per provare a rispondere a queste domande viene dal mondo delle scienze politiche, e in particolare di quella letteratura anglofona che volge lo sguardo alle cosiddette crisi umanitarie.

L'approccio foucaultiano, la cui peculiarità potrebbe forse essere individuata nella capacità di non separare la storia di una disciplina che ambisca alla scientificità dalle forme di esercizio del potere sovrano che la in-forma, innerva fruttuosamente le ricerche delle docenti Katja Lindskov Jacobsen e Larissa Fast, e in particolare la loro analisi congiunta della nozione di *humanitarian governance*, proposta da Michael Barnett nel 2013. Il gesto di Jacobsen e Fast è quello di mettere al centro del dibattito sulle politiche umanitarie l'utilizzo delle tecnologie digitali e dei dati digitali. La domanda di Barnett («who “show up” in the context of humanitarian crisis?»⁵) deve essere necessariamente accompagnata, secondo le autrici, da una domanda intorno alle tecnologie impiegate («what technologies “show up”?»⁶). Non solo quindi, innanzi a una crisi umanitaria, occorre chiedersi chi si prende carico di quei soggetti che, per definizione, hanno una agency limitata, perché le intenzioni e le modalità della presa in carico non siano ingenuamente estinte nella categoria retorica del prendersi cura, ma anche quali tecnologie vengono impiegate e contestualmente sperimentate nel cono d'ombra normativo dettato dallo stato di emergenza. Non si può poi sottovalutare il fatto che la rapidità e la pervasività senza precedenti del progresso tecnologico ha esteso il campo di caccia dalla terra al *Cloud*: ieri si sperimentava sui coloni per offrire un vaccino ai colonizzatori; oggi si sperimenta sui corpi dei migranti per governare nello spazio digitale. Se la posta in gioco delle nuove tecnologie non è mai stata riducibile al qui e ora, la peculiarità delle tecnologie digitali sta nel loro oltrepassare una vera e propria soglia di governamentalità. Il loro utilizzo nel respingimento dei migranti dalla Fortezza Europa avanza di pari passo

⁵ Cfr. M.N. Barnett, *Humanitarian governance*, in «Annual Review of Political Science», 2013, vol 16, n. 1, pp. 379-398.

⁶ K. Lindskov Jacobsen, L. Fast, *Rethinking Access : how humanitarian technology governance blurs control and care*, in «Disasters», 2019, vol. 43, n.s. 2, pp. 151-168: 152.

con una politica di estrattivismo del petrolio digitale, i cui effetti sono tutt'ora sconosciuti: i dati forniti dai soggetti migranti permettono di sperimentare nel laboratorio digitale le politiche del governo a venire.

E se Joseph Paul Jernigan, l'omicida texano reso immortale dal *Visible Human Project*⁷, aveva per lo meno acconsentito a donare il proprio corpo alla scienza, prima di essere giustiziato e tagliato in 1871 fettine, poi diligentemente scannerizzate, l'imperativo impersonale che trionfa dal citato corso foucaultiano (*il faut*) sembra aver definitivamente espunto dal suo meccanismo la parola dei diretti interessati. Bisogna prendersi cura del fuori norma, sottrarlo alla vista se serve, integrarlo se conviene, respingerlo se necessario, ma in ogni caso metterne a profitto il corpo biologico e archiviare quello digitale. E le tecnologie digitali di oggi si prendono cura, in silenzio, del silenzio dei corpi vili. Il primo gesto è quello dell'invisibilizzazione: la comune esperienza di cercare su un motore di ricerca l'acronimo IA si risolve nel suo identificarlo con le parole Intelligenza Artificiale, laddove il suo scioglimento potrebbe rilevare, certo all'interno di una letteratura di altro genere, quello che si scopre essere, come vedremo, il terreno privilegiato del suo esercizio. Le vocali I e A indicano infatti anche il riferimento alle politiche in tema di Immigrazione e Asilo. Il gioco proposto nel titolo, *bIAs*, trova qui il proprio orizzonte di senso. Le tecnologie *IA-based* vengono sperimentate nel campo dell'Immigrazione e delle richieste di Asilo, esercitando una discriminazione già a questo livello, cioè prima dell'appello alla dimensione digitale. In altre parole il primo *bias*, la prima selezione orientata, ha a che vedere con il campione umano sul quale sperimentare le nuove tecnologie. Il report dell'Oxford Department of International Development del 2023, dal titolo *Automating Immigration and Asylum: The Use of New Technologies in Migration and Asylum Governance in Europe*⁸, traccia a favore di lettore un diagramma di intellegibilità del fenomeno: sull'asse delle ordinate, Derya Ozkul, autrice del documento, inserisce i tre momenti del fenomeno migratorio, fotografati come stazioni di sosta dal punto di vista dell'Europa (il paese di provenienza dal quale si pianifica la partenza, il passaggio della frontiera, la richiesta di asilo o di cittadinanza all'interno degli stati membri dell'unione); sull'asse delle ascisse, troviamo invece le specificità delle diverse tecnologie digitali e *IA-based*, che vanno dal riconoscimento vocale alle tecnologie *emotion detection*, dal riconoscimento biometrico allo *speech-to-text*. Alla bidimensionalità cartesiana si aggiungono i rilievi della geografia politica: sullo schema così composto si dispongono allora le specificità delle singole nazioni dell'Unione, in ragione delle loro esigenze e dei vincoli posti da una parte dalla legislazione nazionale, dall'altra da quella che con un termine generale, ma forse non

⁷ Per approfondire la questione: National Library of Medicine, *The Visible Human Project*, link: https://www.nlm.nih.gov/research/visible/visible_human.html, ultima consultazione: 26.05.2025.

⁸ D. Ozkul, *Automating Immigration and Asylum: The Uses of New Technologies in Migration and Asylum Governance in Europe*, Oxford: Refugee Studies Centre, University of Oxford, Oxford 2023.

per questo non adatto, si potrebbe definire la cultura dei singoli paesi (storia, sensibilità, tradizione). Ne emerge una costellazione di casi studio che comprende tecnologie attualmente utilizzate, tecnologie in fase di sperimentazione e progetti finanziati o dismessi dai singoli paesi. Ecco che il contestato e poi abbandonato *iBorderCtrl* (sistema di riconoscimento delle emozioni con focus speciale sulle menzogne), sperimentato alle frontiere in Ungheria, Lettonia e Grecia, viene analizzato vicino al sistema di automatizzazione del processo di richiesta di cittadinanza in Norvegia, completamente automatico, in caso di esito positivo, dall'agosto 2020. Altri casi studio riguardano gli strumenti di traslitterazione dei nomi dei richiedenti asilo nell'alfabeto latino, allo scopo di evitare errori di spelling e uniformare i nomi (se non c'è uno strumento univoco, ogni nazione registra con nomi diversi la stessa persona – es. Mahmood in inglese, Mahmud in tedesco, Mahmoud in francese), che forniscono indirettamente uno strumento di verifica della storia dei migranti ("The [this] name is used [rarely/very rarely] in the indicated country [Syria]. Instead, in [the countries / the country][Libya, Algeria and Morocco], it is used frequently"). Completano il quadro servizi di sorveglianza che utilizzano droni, sensori *offshore* e satellitari per identificare persone e mezzi che viaggiano per i confini europei, l'utilizzo di tecnologie cosiddette *lie detection*, molto controverse e anche molto criticate, passando per *IA-based software* che individuano i furti di identità incrociando documenti (permessi di soggiorno, certificati di nascita, matrimonio e morte), e flirtando con la tecnologia *blockchain*. Tralascieremo in questa sede le legittime domande che riguardano da una parte l'affidabilità di questi strumenti, diversa non solo a seconda del tipo di tecnologia (a oggi lo *speech-to-text* è molto più preciso dell'*emotion detection*), ma anche rispetto alla specificità degli input inseriti (l'utilizzo della traslitterazione come cartina di tornasole per indicare la provenienza dei migranti ha un tasso di successo intono al 90% quando si tratta di nomi che afferiscono al bacino siriano e all'Iraq, ma crolla al 40% nel caso dei nomi arabi, segnalando, nel 34% dei casi, dunque in un caso su tre, presunte incongruenze nelle testimonianze dei migranti) e, dall'altra, la questione della *transparency*, una sorta di correlato non oggettivo della *privacy*. La domanda che porremo affronterà piuttosto la seguente questione: in che modo i corpi dei migranti diventano il laboratorio di sperimentazione di queste nuove tecnologie, in contrapposizione alle indicazioni del regolamento sull'intelligenza artificiale del Parlamento europeo e del Consiglio? Il primo punto del riferimento obbligato in materia dichiara esplicitamente come scopo del regolamento la promozione «di un'intelligenza artificiale (IA) antropocentrica e affidabile»⁹. Si potrebbe forse, senza allontanarsi

⁹ Si riporta per esteso il primo punto del Regolamento (UE) 2024/1689 del Parlamento europeo e del consiglio del 13 giugno 2024, reperibile sulla gazzetta ufficiale dell'Unione europea (link: https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=OJ:L_202401689, ultima consultazione: 26.05.2025), p. 1: «(1) Lo scopo del presente regolamento è migliorare il funzionamento del mercato interno istituendo

troppo dalle intenzioni di chi in primo luogo ne ha coniato l'espressione, immaginare che con intelligenza artificiale antropocentrica, traduzione dell'inglese *human-centred AI*, si intenda rivendicare una certa autonomia del pensiero umano rispetto a quello delle macchine. L'aggettivo afferirebbe dunque all'area semantica del controllo e della verificabilità, se non riconducibile alla sola ragione calcolante, per lo meno da quest'ultima riconoscibile. Proseguendo lungo il regolamento, emerge un elemento paradossale, un rovesciamento umoristico del concetto: l'applicazione di leggi e norme che si rifanno all'antropocentrismo finisce per svuotare il contenuto metafisico del termine *anthropos*, che si traduce, come vedremo, in un principio di inclusione escludente, che autorizza di fatto usi, esplicitamente considerati inaccettabili, delle tecnologie in oggetto su specifici soggetti, privati del titolo di *anthropos*. Andando avanti nella lettura, si apprende infatti che l'intelligenza artificiale consiste in una famiglia di tecnologie in rapida evoluzione che contribuisce al conseguimento di un'ampia gamma di benefici a livello economico, ambientale e sociale nell'intero spettro delle attività industriali e sociali. Si apprende altresì che l'uso dell'IA, in virtù delle sue capacità predittive, può fornire vantaggi competitivi fondamentali alle imprese e condurre a risultati vantaggiosi sul piano sociale e ambientale, dall'agricoltura all'istruzione, dallo sport alla gestione delle infrastrutture. Già al punto cinque però la linearità della favola dell'intelligenza artificiale come panacea dei mali del mondo, genio compensativo dei peccati della Rivoluzione industriale, si increspa:

l'IA può nel contempo, a seconda delle circostanze relative alla sua applicazione, al suo utilizzo e al suo livello di sviluppo tecnologico specifici, comportare rischi e pregiudicare gli interessi pubblici e i diritti fondamentali tutelati dal diritto dell'Unione. Tale pregiudizio può essere sia materiale sia immateriale, compreso il pregiudizio fisico, psicologico, sociale o economico¹⁰.

Da questo momento in poi, la capacità inferenziale dei sistemi di IA viene convocata in tribunale e chiamata a rendere conto degli strumenti impiegati in

un quadro giuridico uniforme in particolare per quanto riguarda lo sviluppo, l'immissione sul mercato, la messa in servizio e l'uso di sistemi di intelligenza artificiale (sistemi di IA) nell'Unione, in conformità dei valori dell'Unione, promuovere la diffusione di un'intelligenza artificiale (IA) antropocentrica e affidabile, garantendo nel contempo un livello elevato di protezione della salute, della sicurezza e dei diritti fondamentali sanciti dalla Carta dei diritti fondamentali dell'Unione europea («Carta»), compresi la democrazia, lo Stato di diritto e la protezione dell'ambiente, proteggere contro gli effetti nocivi dei sistemi di IA nell'Unione, nonché promuovere l'innovazione. Il presente regolamento garantisce la libera circolazione transfrontaliera di beni e servizi basati sull'IA, impedendo così agli Stati membri di imporre restrizioni allo sviluppo, alla commercializzazione e all'uso di sistemi di IA, salvo espressa autorizzazione del presente regolamento».

¹⁰ Id., p. 2.

materia di “identificazione biometrica”, di “categorizzazione biometrica”, nonché di “riconoscimento delle emozioni”. Si tratta però di un processo in sordina, perché al posto della sentenza troviamo un regolamento suonato in chiave condizionale. Così recita il punto 28):

L’IA presenta, accanto a molti utilizzi benefici, la possibilità di essere utilizzata impropriamente e di fornire strumenti nuovi e potenti per pratiche di manipolazione, sfruttamento e controllo sociale. Tali pratiche sono particolarmente dannose e abusive e dovrebbero essere vietate poiché sono contrarie ai valori dell’Unione relativi al rispetto della dignità umana, alla libertà, all’uguaglianza, alla democrazia e allo Stato di diritto e ai diritti fondamentali sanciti dalla Carta, compresi il diritto alla non discriminazione, alla protezione dei dati e alla vita privata e i diritti dei minori¹¹.

Si articola così un sistema di *soft law*, per usare l’espressione utilizzata da Teresa Numerico, perché in larga misura si interviene con delle raccomandazioni nelle quali si fa fatica a identificare obblighi precisi, con rare eccezioni, prevalentemente di carattere burocratico. Si crea così una architettura di riferimento internazionale in cui l’adozione di sistemi *IA-based* da parte degli stati membri deve essere notificata entro 30 giorni dalla sua adozione e porta con sé i famosi obblighi di trasparenza, estinti, il più delle volte, nell’ingiunzione a informare l’utente quando interagisce con un *chatbox*. Come una sorta di tic nervoso, la locuzione “diritto alla non discriminazione” emerge a singhiozzi lungo il regolamento (27, 28, 31, 45, 48, 56, 57, 58, 60, 67, 70), e nel punto 60 trova tematizzazione nel settore della migrazione, dell’asilo e della gestione del controllo delle frontiere, in una formulazione linguistica urgente e insieme fragile:

I sistemi di IA nel settore della migrazione, dell’asilo e della gestione del controllo delle frontiere non dovrebbero in alcun caso essere utilizzati dagli Stati membri o dalle istituzioni, dagli organi o dagli organismi dell’Unione come mezzo per eludere gli obblighi internazionali a essi derivanti a titolo della convenzione delle Nazioni Unite relativa allo status dei rifugiati firmata a Ginevra il 28 luglio 1951, modificata dal protocollo del 31 gennaio 1967. Essi non dovrebbero essere utilizzati per violare in alcun modo il principio di non respingimento o per negare sicure ed efficaci vie legali di ingresso nel territorio dell’Unione, compreso il diritto alla protezione internazionale¹².

A indebolire ulteriormente il quadro normativo concorre l’annesso III, ossia la sezione che prevede delle deroghe rispetto all’uso di sistemi prima ritenuti a rischio

¹¹ Id., p. 8.

¹² Id., p. 17.

inaccettabile. Due momenti di questo annesso coinvolgono il migrante: già nel punto 1, l'impiego della biometria viene consentita nei casi in cui occorre confermare che una persona fisica è la persona che dice di essere; ma è al punto 7 che l'annesso tratta esplicitamente di «migrazione, asilo e gestione del controllo delle frontiere»¹³, dove un rischio per la sicurezza, un rischio di migrazione irregolare o un rischio per la salute posto da una persona fisica che intende entrare o è entrata nel territorio di uno Stato membro vengono equiparati come condizione sufficiente alla deroga nell'impiego di sistemi di IA a rischio inaccettabile. L'impressione che se ne trae è allora che i richiedenti asilo e i migranti costituiscano uno spazio di sperimentazione, un terreno liminale all'interno del quale si assiste a una deroga al principio di uguaglianza. La legge disvela un principio di inclusione escludente al cuore dell'idea stessa dell'antropocentrismo che mal si cela dietro la presunta responsabilità di una intelligenza altra rispetto a quella umana, o *human-centred*, che dir si voglia. In altre parole, la legge che dovrebbe vietare la discriminazione attraverso l'uso degli algoritmi, in nome dei principi della Carta, si fa teatro di un rovesciamento comico dei principi ai quali si appella, e che pone la discriminazione all'inizio del proprio esercizio. In principio era il meticcio, sul corpo del quale il verbo si fa legge.

Chi raccoglie i frutti dell'innovazione tecnica? Quali corpi ne pagano il prezzo? In questo limbo normativo, in cui si apre lo spazio della sperimentazione tecnica, si insinua altresì il sospetto che il meticcio epistemico sul quale lo sperimentale agisce debba la propria condizione a una discrasia in grado di abbattersi sul biologico a partire da pregiudizi di natura culturale e sociale, la cui invariante non sta nell'oggetto discriminato, bensì nelle modalità del suo imbrigliamento. La questione dei *bias* e delle discriminazioni potrebbe essere allora pensata nella sua componente soggettiva, squisitamente umana, come la necessità, elevata a metodo dalla ragione scientifica, di individuare corpi da dare in pasto all'evoluzione tecnica. Questo approccio ha, se non altro, il vantaggio di non dispensare nessuno dalla presa di discorso intorno al tema, sollecitando l'intelligenza collettiva (umana e non) per una distribuzione equa dei diritti e delle responsabilità. In questo senso, non sarebbe forse del tutto scorretto e infruttuoso pensare in maniera rigorosa il momento presente come una soglia di trasformazione, una soglia epistemologica di governamentalità, a partire dalla constatazione della quale è possibile separare l'ideologia scientifica, quella di una tecnogenesi che esclude l'*anthropos* e dalla quale occorre difendersi o alla quale è necessario abbandonarsi, dalla necessità di comprendere le conseguenze e le potenzialità di una tecnologia alla quale richiediamo *transparency*, ma il cui primo effetto è quello di gettare una luce su una opacità nel cuore dell'antropogenesi.

¹³ Id., p. 128.

