

Elettra Stradella

Professoressa associata di Diritto pubblico comparato, Università degli Studi di Pisa. Mail: elettra.stradella@unipi.it

Credo che per rispondere a questi quesiti occorra partire in primo luogo dallo stato dell'arte tecnologico, e questo mostra il crescente e soddisfacente impiego dell'AI in ambito biomedico: dalle svariate applicazioni in ambito diagnostico (interpretazione dei tracciati elettrocardiografici, previsione di eventi correlati a patologie quali il diabete, o malattie autoimmuni), fino all'elaborazione di paradigmi Brain-Computer Interface (BCI) per l'interpretazione dell'attività cerebrale, settore ancora prevalentemente appannaggio della ricerca più che dell'applicazione medica.

Le condizioni di invecchiamento generale della popolazione, da cui deriva la necessità di confrontarsi con un ampio numero di persone anziane e non-autosufficienti, in molti casi affette da patologie croniche, rendono particolarmente utile il potenziamento di modelli diagnostici e di sistemi di monitoraggio, che in parte già beneficiano, e di più beneficerebbero, degli enormi vantaggi computazionali dell'AI, incidendo (anche) sulla sostenibilità dei sistemi sanitari.

D'altra parte, credo che sia essenziale, prima di concentrarmi sulla questione del rapporto medico-paziente, dell'autonomia della professione medica, e sulle trasformazioni del principio del consenso informato, sottolineare una prima criticità che l'utilizzo dell'AI in ambito sanitario pone, e che è più in generale rilevabile nei sistemi di intelligenza artificiale. In particolare, la presenza di bias derivanti dai dati messi a disposizione dei sistemi, e dunque i fenomeni di amplificazione delle esistenti discriminazioni, rischiano di produrre in ambito sanitario effetti di penalizzazione dei soggetti in relazione a caratteristiche personali ad alto grado di vulnerabili-

tà e stereotipizzazione, quali l'identità sessuale o di genere, l'origine etnica, l'appartenenza religiosa.

Ci sono esempi concreti di questo fenomeno: basti pensare agli algoritmi di previsione utilizzati negli Stati Uniti per l'individuazione di pazienti con esigenze sanitarie complesse (Obermeyer, 2019). In questo caso è stato recentemente mostrato come a un determinato livello di rischio i pazienti neri stiano peggio, siano più malati dei pazienti bianchi, e risolvere questa disegualianza sociale e assistenziale comporterebbe un aumento dei pazienti neri che ricevono ulteriori cure fino a quasi il 50%. Ma perché l'algoritmo utilizzato per scegliere verso quali pazienti destinare la spesa sanitaria, in questo caso, produce in maniera così evidente una disparità? Perché automaticamente prevede, e dunque richiede, minori costi per l'assistenza sanitaria per i pazienti neri rispetto ai pazienti bianchi. È probabile che questo derivi, ancora una volta, da una storia di disparità e discriminazioni: la tendenziale minore attenzione alla salute dei pazienti neri, sottoposti ad un'intensità di cura inferiore a fronte di sintomatologie o patologie analoghe o corrispondenti a quelle dei pazienti bianchi, si traduce in un algoritmo che paradossalmente, considerandoli meno bisognosi di cure, li condanna a peggiori condizioni di salute.

D'altra parte, le questioni relative alla qualità dei dati, e alla presenza di bias, sembrano assumere sempre maggiore rilevanza, anche nella riflessione dei tecnologi, ragione per cui non sembra irrealistico ritenere che nei prossimi anni sarà possibile progettare sistemi antidiscriminatori *by default* e *by design*, in grado di superare le criticità delineate.

Una volta limitate o risolte queste criticità, non resteranno dubbi sull'importante funzione predittiva dell'AI in ambito sanitario. Da questo

punto di vista, la capacità predittiva dei sistemi di AI sarà sempre più essenziale, a mio avviso, quando sarà utilizzata per una funzione che definirei: *ingegneristica, sistemica, epidemiologica*.

Basti pensare alla vicenda su cui si concentra in questi giorni tutta la nostra attenzione, il dilagare dell'epidemia di Covid-19, e all'impatto che su fenomeni del genere potrebbe avere l'utilizzo dell'Intelligenza Artificiale.

Come è stato sottolineato da tecnologi e operatori dell'innovazione (Lambrecht, *Global Health and Life Sciences Practice – SAS*) l'AI può agire su diversi versanti: dalla rilevazione dei primi sintomi indicativi di una nuova epidemia prima che questa sia conclamata, con gli ovvi effetti di limitazione della diffusione del virus responsabile (sintomi che non sono necessariamente clinici), al completamento per i risultati clinici, le specifiche complicanze e le caratteristiche di una nuova epidemia, essenziale per consentire al sistema sanitario di rispondere con efficacia, fino all'utilizzo in attività di automazione rivolte a medici e cittadinanza (ad esempio *chatbot* per la segnalazione di sintomi, con capacità di raccolta di informazioni immensamente superiori a quelle che possono mettere in campo gli strumenti tradizionali quali "numeri verdi" o simili). Un simile utilizzo dell'AI in sanità, ed anche, in senso lato, nella ridefinizione del rapporto medico-paziente può condurre a risultati virtuosi quando inserito all'interno di un generale sistema di e-health.

Il discorso cambia quando da un utilizzo "ingegneristico", "sistemico" e "epidemiologico" si passa alla configurazione dell'AI come vera e propria "assistente cognitiva", che affianca il medico nelle attività diagnostiche e di cura.

Esempi già ne esistono: basti pensare a Watson, di IBM, programma di intelligenza artificiale rivolto agli oncologi per la definizione di cure (o-

sannato nel 2017 per la sua capacità di diagnosticare un tumore in dieci minuti, e vituperato un anno dopo in un'inchiesta della rivista *Stat* che ne metteva in luce la notevole fallibilità derivante, probabilmente, ancora una volta dai dati, perché le sue decisioni sarebbero state assunte sulla base di dati riguardanti un numero limitato di medici e pochi ospedali), o all'algorithm "Medical Sieve": un sistema di "assistenza cognitiva automatizzata" dedicata a radiologi e cardiologi, disegnato per aiutarli proprio nella loro attività diagnostica e decisionale. Il sistema, raccogliendo dati clinici, documentali e derivanti da tecniche di *imaging* dei pazienti, li analizza rintracciando, se ve ne sono, anomalie. I risultati della sua ricerca alimentano un algoritmo che utilizza elementi provenienti sia da risultanze cliniche indipendenti del singolo paziente, sia da statistiche relative a pazienti simili, ovviamente nella misura resa possibile dalla potenza computazionale del sistema, per giungere ad una potenziale diagnosi differenziale che orienti il medico nella sua decisione.

Ma un recente stato dell'arte pubblicato su *Nature* mostra la molteplicità delle possibili applicazioni assistenziali e diagnostiche, tra cui il progetto di diagnosi umana *Human Dx*, strumento decisionale basato sull'AI che si avvale dei Big Data, accumulabili attraverso il *crowdsourcing* che coinvolge a livello globale medici e ricercatori. Il progetto ha ricevuto centinaia di migliaia di contributi da decine di migliaia di professionisti distribuiti in oltre 80 Paesi: l'idea è che l'intelligenza collettiva, potenziata dalla capacità computazionale che solo l'AI può mettere in campo, possa condurre ad un successo diagnostico incomparabile rispetto a quello conseguibile dai singoli medici posti di fronte al caso. (McCall, 2020).

È di queste ore la notizia di algoritmi (Alibaba) sviluppati proprio per l'individuazione

dell'infezione da Covid-19 in grado di diagnosticare in 20 secondi la malattia, grazie a sistemi di AI alimentati dai dati campione di più di 5.000 casi confermati di infezione: l'algoritmo batterebbe sul tempo dunque il medico, più lento nell'interpretare un'indagine tomografica e comunque bisognoso di numerose immagini.

Si potrebbero fare molti altri esempi, ma bastano questi, nello spazio a disposizione, per comprendere che l'interazione tra medico e AI pone una serie di problemi che riguardano in particolare la responsabilità del primo, da diversi punti di vista: quello privatistico della responsabilità derivante da danni provocati nei confronti di terzi; quello dell'*accountability* del medico, a cui si collega il rischio della deresponsabilizzazione che, in questo ambito, potrebbe tradursi in un appiattimento decisionale e diagnostico; e infine quello della responsabilità nella relazione di cura, da cui deriva il valore attribuito all'autodeterminazione del paziente.

Il primo punto di vista è quello più "tradizionale", sempre toccato quando ci si interroga sull'utilizzo della robotica e dell'AI, nei vari settori. Questo non significa che sia di immediata soluzione. Com'è noto la dottrina, con particolare riferimento all'ambito medico, utilizza l'art. 2050 del codice civile, cioè la fattispecie della responsabilità per l'esercizio di attività pericolose (Topi, 2016; Parziale, 2019), che pure potrebbe non ritenersi del tutto adeguata a descrivere le peculiarità dell'AI, la cui "pericolosità" non deriva certamente dall'oggetto dell'attività che, in questo caso, altro non sarebbe che la stessa attività diagnostica a cui è chiamato il medico, ma semmai dalle modalità con cui quell'attività è realizzata (statistiche, probabilistiche, matematiche, basate su dati le cui modalità di raccolta ed elaborazione non sempre sono chiare e definite, sia che siano *supervised* sia che siano *unsupervised*, ciò da cui

deriva il noto problema della c.d. *black box*). D'altra parte, nelle modalità sta la natura dei sistemi di AI, che sarebbero così considerati intrinsecamente pericolosi. D'altra parte, anche il richiamo all'opportunità di distinguere, nell'identificazione della responsabilità, tra attività concretamente esercitata dal medico e funzionamento del sistema, pone una serie di dubbi, derivanti in primo luogo dallo stretto legame che si instaura tra i due momenti quando l'AI abbia una funzione di "assistenza cognitiva", come si diceva sopra. È immaginabile di poter operare, magari in sede giudiziale, una quantificazione/qualificazione del contributo effettivamente prestato da parte del medico e, per converso, di quello derivante dalla "macchina"? Ha senso chiedersi se l'errore diagnostico sia derivato da un cattivo funzionamento dell'algoritmo, dai dati che ne hanno in parte determinato il risultato, dalla lettura che il medico ha dato del risultato e dall'applicazione che ha inteso farne? Ha senso immaginare di applicare all'AI categorie giuridiche come quelle dell'"oggetto difettoso"? Credo che le risposte a questi interrogativi debbano essere negative.

Più logica mi sembra invece la considerazione per cui l'individuazione della responsabilità non sarebbe in fondo diversa da quella che avviene per qualsiasi tipo di decisione medica supportata da strumenti, informazioni, dati, che siano esterni all'autonoma capacità e percezione del medico. Dagli strumenti di diagnostica per immagini fino alle analisi del sangue, il medico si avvale abitualmente di qualcosa che è "altro da sé", ma non per questo è messa in discussione la sua responsabilità diagnostica e di cura; lo stesso potrebbe accadere con l'Intelligenza Artificiale.

Il problema però qui è un altro, a mio avviso, e sta nel fatto che con riferimento all'AI risulta particolarmente difficile per il medico verificare

il corretto funzionamento del dispositivo di Intelligenza Artificiale, e dunque il vero rischio sembra essere quello, avanzato anche per l'utilizzo dell'AI nella decisione giudiziaria, di una deresponsabilizzazione di fatto del professionista, che inevitabilmente ricalcherà nelle sue valutazioni i risultati provenienti dal dispositivo intelligente, per la difficoltà di discostarsi da qualcosa che non è in grado di mettere in discussione razionalmente e fondatamente. Anche in questo caso allora, come nelle applicazioni giudiziarie, assume un ruolo centrale quel "diritto alla spiegazione" che si configura come diritto all'intervento umano come diritto in quanto tale, quindi sorretto da una reale comprensione della natura della decisione, medica in questo caso, dei processi che ne stanno a fondamento, del livello di automatizzazione, dei dati che ne hanno orientato il contenuto. Sappiamo che oggi molti *data scientists* stanno lavorando per aprire la *black box*, ma sicuramente il percorso non è compiuto.

La questione allora non riguarderebbe tanto o solo la responsabilità per danni causati nei confronti di terzi, ma la ridefinizione della *soggettività* stessa del medico.

Sotto questo profilo, pur con sfaccettature diverse, la questione della "*spersonalizzazione*" del rapporto medico-paziente si avvicina per certi versi a quella del rapporto giudice-imputato, e non riguarda soltanto la tematica, prettamente civilistica, della responsabilità del medico, rispetto alla quale non sono inedite infatti le fattispecie della responsabilità da "contatto sociale", o il c.d. "contratto di spedalità", in cui il rapporto tra medico e paziente sembra concretizzarsi in un obbligo di prestazione sempre e comunque destinata, se esercitata a regola d'arte, a condurre verso la guarigione, come se l'eventualità della morte non venisse, in fondo, neppure presa in considerazione. Riguarda

soprattutto *la relazione*, e come il dispiegarsi della relazione è in grado di incidere sul diritto all'autodeterminazione dei pazienti.

È interessante, da questo punto di vista, la recente sentenza della Cassazione n. 10423 del 2019, in cui, a fronte della richiesta del ricorrente volta ad ottenere un ristoro non soltanto per il danno alla salute conseguente alla mancata informazione, ma per quello determinato dalla «lesione del diritto all'autodeterminazione terapeutica in sé considerato, rispetto al quale il carattere necessitato dell'intervento e la sua corretta esecuzione restano circostanze prive di rilievo», la Corte sottolinea, riprendendo i suoi precedenti, che «in tema di attività medico-chirurgica, è risarcibile il danno cagionato dalla mancata acquisizione del consenso informato del paziente in ordine all'esecuzione di un intervento chirurgico, ancorché esso apparisse, *ex ante*, necessitato sul piano terapeutico e sia pure risultato, *ex post*, integralmente risolutivo della patologia lamentata, integrando comunque tale omissione dell'informazione una privazione della libertà di autodeterminazione del paziente circa la sua persona, in quanto preclusiva della possibilità di esercitare tutte le opzioni relative all'espletamento dell'atto medico e di beneficiare della conseguente diminuzione della sofferenza psichica, senza che detti pregiudizi vengano in alcun modo compensati dall'esito favorevole dell'intervento». E com'è noto quando la Corte costituzionale, nel 2008, afferma la fundamentalità del diritto al consenso informato, ne definisce il contenuto nella «sintesi di due diritti fondamentali della persona: quello all'autodeterminazione e quello alla salute». Dunque se l'attività del medico trova il suo fondamento nel consenso, e il consenso, per come costruito dalla giurisprudenza, è chiaramente definito attraverso la relazione che si instaura tra medico e paziente, non risolvendosi in una

mera attività informativa o un adempimento burocratico, ma il consenso non soltanto è funzionale all'attuazione del diritto all'autodeterminazione, ma quest'ultimo ne definisce il contenuto stesso, possiamo dire che la professione medica abbia una specificità che sta nelle modalità con cui può essere misurato il risultato dell'azione finalizzata alla tutela del diritto alla salute. Come scrive la Cassazione l'assenza del consenso informato rende «irrilevante il fatto che l'intervento medico sia stato eseguito in modo tecnicamente e oggettivamente corretto».

Il pregiudizio arrecato al paziente non informato, non ascoltato, limitato nella sua autodeterminazione, non troverebbe compensazione, in questa visione, nell'esito favorevole dell'intervento chirurgico, nel caso di specie, in generale nella cura. Ciò significa che il compito del medico non è soltanto quello di compiere un'attività diagnostica, di prendere delle decisioni di cura ed applicarle, e, in fin dei conti, non è neppure soltanto quello di guarire, ma si sviluppa, come detto, nella dimensione relazionale.