

UNIVERSITÀ DEGLI STUDI DI SALERNO

DIPARTIMENTO DI STUDI POLITICI E SOCIALI



**DOTTORATO DI RICERCA IN
“SOCIAL THEORY, DIGITAL INNOVATION AND PUBLIC
POLICIES”**

CICLO XXXV

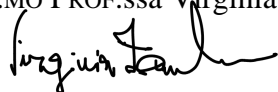
TESI DI DOTTORATO IN:

“Sfide etiche e legali dell'assistenza sanitaria basata sull'intelligenza artificiale: la responsabilità del medico e dell'amministrazione sanitaria”

COORDINATORE

CHIAR.MA PROF.SSA  Giuseppina Cersosimo

TUTOR

CHIAR.MO PROF.SSA  Virginia Zambrano

CANDIDATO

DOTT. SSA  Martina Itri

ANNO ACCADEMICO

2021/22

CAPITOLO I

Profili introduttivi dell'intelligenza artificiale in medicina

1. Introduzione.....pag. 1
2. Intelligenza artificiale e innovazione tecnologica.....pag. 4
3. Decisioni algoritmiche e principali sistemi di apprendimento.....pag. 7
4. I Big Data e la medicina.....pag. 13
5. Il fenomeno dei black box applicato al settore sanitario.....pag. 17
6. Sistemi robotici e l'agente intelligente in sanità.....pag. 25

1. Introduzione

La responsabilità medica rappresenta un ramo importante del diritto nel quale, in primo luogo, si confrontano le aspettative di guarigione dei pazienti e il dovere del medico di fornire loro un'assistenza adeguata e possibilmente risolutiva.

Purtroppo la medicina non è una scienza esatta, accade quindi che il paziente possa non essere sempre pienamente soddisfatto della cura che ha ricevuto: è proprio in questo caso che entra in gioco la legge, con il primo obiettivo di prevenire le situazioni di pericolo e il secondo obiettivo di garantire il risarcimento del danno cagionato da quel medico che non ha agito con diligenza, prudenza, competenza, e dall'azienda sanitaria che ha disatteso gli obblighi contrattuali conclusi con il paziente.

Oggi, una delle principali questioni che animano il dibattito attorno alla responsabilità medica è quella relativa all'uso degli strumenti di Intelligenza Artificiale in medicina, si tratta di elementi terzi ingegneristici, biomedici, di natura meccanica o informatica, che supportano il lavoro quotidiano del medico.

Nel nostro ordinamento l'individuazione del responsabile diventa ancora più complessa se si pensa alle modifiche legislative che sono intervenute nella materia sanitaria, infatti si è assistito ad un passaggio dalla responsabilità contrattuale per il medico alla responsabilità, ricondotta da dottrina e giurisprudenza, nell'alveo della responsabilità extracontrattuale.

Nel sistema sanitario il ricorso alle intelligenze artificiali è improntato al virtuoso risultato.

Questo tipo di tecnologia è in grado di dare una percentuale di rischio a sviluppare, in base a determinati risultati, determinate malattie, facendo ricorso ad alcune caratteristiche dei pazienti già ricoverati.

E' opportuno ricordare che l'attività medica si basa sull'interpretazione della narrazione dell'uomo, dell'esame obiettivo e delle valutazioni cliniche, terapeutiche e diagnostiche del suo corpo.

L'anamnesi costituisce il vero fulcro della prestazione sanitaria in quanto finalizzata alla raccolta di informazioni personali atte ad indagare il malessere del paziente.

Il problema nasce dal fatto che la scienza medica non è basata solo sull'evidenza scientifica ma anche sulla specificità del caso e sulla narrazione del paziente.

L'interazione uomo-macchina nelle azioni/omissioni pone problemi di responsabilità in caso di danni a terzi.

La responsabilità, infatti, in questi casi non è di facile soluzione. Il problema è dato dal fatto che questo nuovo approccio tecnologico porta con sé realtà nuove che difficilmente riescono a trovare un riferimento normativo, anche in via analogica, nel codice civile.

Il punto di partenza di questo studio di tesi è dato dalla dicotomia che si è creata tra intelligenza artificiale e intelligenza umana. All'esito delle ricerche fatte si capirà che si tratta di una falsa contrapposizione.

Chi pensa che un algoritmo possa scavalcare il parere di un medico, è in errore. Anche la tecnologia è fallibile e la sua supervisione sarà necessariamente affidata all'uomo.

Per poter essere di aiuto in ambito sanitario un algoritmo di AI deve garantire prestazioni pari o superiori a quelle di un essere umano nei compiti specifici che gli vengono assegnati.

Ma molte delle decisioni che i medici prendono quotidianamente, sono incredibilmente complesse e richiedono un approccio ben più articolato rispetto a quello offerto dall'utilizzo dell'intelligenza artificiale nell'elaborazione dei dati.

Si potrebbe far riferimento "all' autonomia dell'intelligenza artificiale" cercando di individuare il quantum di "apporto" dell'essere umano: capire quando è minimo e quindi l'intelligenza artificiale è del tutto autonoma anche nei suoi aggiornamenti oppure quando l'intervento dell'uomo è imprescindibile.

Nell'ambito degli enti ospedalieri, però, l'intelligenza artificiale, ha quasi sempre necessità del contributo umano perché è attraverso quel contributo che si arriva al risultato definitivo. Ecco che allora sorgerà il bisogno, al fine di non addebitare responsabilità inesistenti, di analizzare la responsabilità, se del caso scindendone i vari profili.

Per realizzare l'enorme potenziale dell'IA per trasformare l'assistenza sanitaria in meglio, gli stakeholder nel campo dell'IA, compresi i responsabili dell'IA, i medici, i pazienti, gli etici e i legislatori, devono essere impegnati nel dibattito etico e legale su come l'IA viene implementata con successo nella pratica.

La ricostruzione della responsabilità che questo studio di tesi si impegna a realizzare procederà tramite orientamenti giurisprudenziali e dottrinali e la ricostruzione delle strategie riguardanti l'etica e il diritto dell'IA nella sanità nella Comunità Europea.

Per quanto si parli in maniera omnicomprensiva di sistemi di intelligenza artificiale, si tratta poi nel concreto di sistemi o prodotti che presentano notevoli differenze. Basti pensare al fatto che il robot chirurgico viene utilizzato da un soggetto che ha una specifica competenza e formazione professionale, mentre l'automobile driverless potrebbe in astratto essere guidata anche da una persona con ridotte capacità sensoriali, per rendersi conto delle differenze sussistenti nel campo delle intelligenze artificiali.

In questo senso, occorre chiedersi fino a che punto una soluzione normativa che abbracci tutte le tipologie di intelligenza artificiale, senza operare specifiche differenziazioni in ragione delle loro precipue caratteristiche e del settore ove impattano, rappresenti la soluzione preferibile.

Attualmente non esiste un quadro normativo dell'UE completamente armonizzato per la responsabilità in materia di IA e robotica come cura e robot medici. Tuttavia, l'Europa ha compiuto diversi passi per affrontare la questione dell'etica e della responsabilità.

E' complesso stabilire se le proposte formulate in seno all'Unione Europea siano in grado di rispondere adeguatamente alla molteplicità dei problemi determinati dalla robotica avanzata. Esse hanno sicuramente il merito di prevedere l'adozione di regole giuridiche nuove per le intelligenze artificiali.

Si tratta di un percorso essenziale al fine di fornire certezza tanto agli operatori del settore quanto agli utilizzatori di smart products e sistemi avanzati di intelligenza artificiale, con l'introduzione di un regime di responsabilità oggettiva in capo a quei soggetti che sembrano meglio in grado di contenere i rischi ed evitarne o diminuirne le conseguenze negative.

2. Intelligenza artificiale e innovazione tecnologica

Il termine "Intelligenza artificiale" o nella forma abbreviata "IA" è ampiamente diffuso nella società odierna e ricomprende diverse branche delle competenze e delle capacità produttive dell'uomo.

L'intelligenza artificiale è correlata al progresso tecnologico e ricomprende le attività che normalmente richiedono l'intervento dell'uomo, come la percezione visiva, il riconoscimento del linguaggio, prendere delle decisioni, diagnosticare una prognosi.

Oggi giorno stiamo assistendo ad una rivoluzione culturale, tecnologica e scientifica incentrata sull'applicazione dell'apprendimento automatico a database di grandissime dimensioni, che permettono alle macchine di analizzare grandi quantità di dati più velocemente e meglio degli esseri umani.¹

Appare difficile poter dare una definizione univoca di "IA" articolandosi quest'ultima in una serie di sottotipi come il "machine learning", il "deep learning" "algoritmi" e "big data".

¹ N. Musacchio, G. Guaita, A. Ozzello, M.A. Pellegrini, P. Ponzani, R. Zilich, A. De Micheli "Intelligenza Artificiale e Big Data in ambito medico: prospettive, opportunità, criticità." in The Journal of AMD, 2018, Vol. 21-3, pag. 204

Originariamente il termine “intelligenza artificiale” è stato coniato da John McCarthy, informatico statunitense, che l’ha definita come “*science and engineering of making intelligent machines*”².

L’ I.A. combina varie tecniche diverse, come sono diversi i sistemi di elaborazione del linguaggio. Nel linguaggio comune queste tecniche sono spesso indicate come “algoritmi”, un costrutto matematico che descrive una “*struttura di controllo che realizza uno scopo specifico con determinati input*”.

Fin dalla rivoluzione industriale, si è verificato un vasto sviluppo nel campo della tecnologia che ha sostituito il lavoro manuale in vari campi.

L’ IA può essere definita come una branca della scienza che crea macchine intelligenti e programmi per computer per eseguire vari compiti che richiedono l’intelligenza umana. In sintesi è un sistema che imita varie funzioni che un essere umano può svolgere.

Un tempo l’IA era solo un concetto fantascientifico e un argomento dei dibattiti in cui si discuteva dell’effetto della tecnologia nel mondo moderno, ma oggi è diventata parte integrante della nostra vita quotidiana. Essa è diventata la funzione chiave di molti settori tecnici e vari, come produzione, assistenza sanitaria, catene di approvvigionamento, ecc

La capacità dell’IA di fare cose che l’uomo non può, si traduce in miglioramento delle prestazioni e della produttività.³

Il salto di qualità dei sistemi di intelligenza artificiale è avvenuto quando l’attenzione si è spostata sulla possibilità di utilizzare i metodi di apprendimento automatico (machine learning) a grandi masse di dati.

L’uomo, anziché fornire tutta la conoscenza alla macchina, le fornisce un metodo di apprendimento, da applicare ai dati a cui la macchina ha accesso, per estrarre le indicazioni su come svolgere il compito affidatole.

La scienza moderna che studia l’intelligenza artificiale parte dal pensiero elaborato di Alan Turing, che si chiese come si potesse dimostrare la differenza tra umani e computer, per quanto riguarda l’intelligenza.

Il matematico britannico propose un gioco che può essere utilizzato per rispondere alla domanda “Può la macchina pensare?” Questo gioco di imitazione è noto appunto come test di Turing.

² McCarthy J. “A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence” in AI Magazine, 2006, vol. 27, ser. 4.

³ Karthikeyan J., Ng Yu Jin, E Ting Su Hie “*Learning Outcomes of Classroom Research*, New Delhi: L’Ordine Nuovo Publication, 2022.

Il gioco si articola partendo da tre persone, un uomo (A), una donna (B) e un interrogatore (C) che può essere di entrambi i sessi. L'interrogante rimane in una stanza separata dagli altri due.

Il test di Turing identifica se un computer ha un'intelligenza artificiale tale che può, in determinate condizioni, imitare gli umani in modo ingannevolmente realistico.

Durante una conversazione con un interlocutore durante il test, l'umano non dovrebbe più distinguere chiaramente una persona da un computer.

Il test è superato quando la persona testata non è più in grado di distinguere tra le risposte date dal computer e quelle date da una persona reale.

Per superare il test di Turing, un computer avrebbe bisogno almeno delle seguenti competenze:

- Elaborazione del linguaggio naturale: deve essere in grado di comunicare in modo naturale in una lingua come l'inglese.
- Rappresentazione della conoscenza: deve essere in grado di avere conoscenza e conservarla in qualche luogo.
- Ragionamento automatizzato: deve essere in grado di ragionare in base a quanto memorizzato.
- Machine learning: deve essere in grado di imparare dall'esperienza.⁴

Secondo alcuni critici, il test dovrebbe essere classificato come controverso. Ci sono aree in cui un computer è bravo a rispondere alle domande e aree in cui le risposte potrebbero essere ingannevoli, come ad esempio alle domande alle quali è possibile rispondere con un sì o con un no.

Naturalmente non va dimenticato che il test di Turing è stato sviluppato a metà del secolo scorso e nonostante sia un utile punto di partenza per comprendere il complesso meccanismo dell'intelligenza artificiale, appare ormai superato. Infatti sono stati creati altri tipi di test che, ad esempio, elaborano vari contenuti multimediali con l'aiuto dell'intelligenza artificiale, della musica, del parlato, delle immagini o dei video.⁵

Se è chiaro da questo breve excursus che la definizione principale di "intelligenza artificiale" viene fornita dal mondo scientifico, è vero che anche gli studiosi del diritto si sono cimentati nel tentativo di fornire una definizione di "IA".

⁴ Roncaglia G. *“Computer Che Copiano: Test Di Turing, Web Corpora e Filtraggio Collaborativo”*, 2014, si veda <http://hdl.handle.net/2067/2614>.

⁵ Ribola, J. *"Il test di Turing è ancora rilevante? Problemi e attualità di un'icona dell'Intelligenza Artificiale"* Dipartimento di Comunicazione ed Economia, Università degli Studi di Modena e Reggio Emilia, 2016

Ad esempio la definizione suggerita di recente dalla Commissione Europea si riferisce all'IA come "sistemi che mostrano comportamenti intelligenti analizzando il loro ambiente e intraprendendo azioni - con un certo grado di autonomia - per ottenere risultati specifici obiettivi"⁶."

Sulla base di quanto elaborato dal gruppo di esperti istituito dalla Commissione, i sistemi di intelligenza artificiale, sono sistemi software o hardware progettati dagli esseri umani che, individuato un obiettivo, acquisiscono, elaborano e interpretano dati, ragionando sulla base delle informazioni e delle conoscenze immagazzinate.⁷

Infatti, molte tecnologie di IA richiedono una moltitudine di dati per migliorare le loro prestazioni, e quando raggiungono un buon livello di prestazione, esse possono contribuire a migliorare ad automatizzare il processo decisionale.

E' chiaro che si è davanti ad una rivoluzione culturale, tecnologica e scientifica. Una scienza che "simula il ragionamento umano" non può che sollevare dubbi sul fatto che questo fenomeno sia sotto il controllo umano.⁸

Il presente lavoro di tesi tende a fornire una sintesi sulle tecnologie di intelligenza artificiale e sull'impatto nella vita quotidiana degli esseri umani.

3. Decisioni algoritmiche e principali sistemi di apprendimento

La locuzione "decisioni algoritmiche" ha carattere generale poiché individua una serie di istruzioni messe in sequenza in modo tale da definire un'attività che può riguardare diversi campi, attraverso calcoli matematici.

⁶ Si veda "Comunicazione della commissione al Parlamento europeo, al Consiglio europeo, al Consiglio, al Comitato economico e sociale europeo e al comitato delle regioni - piano coordinato sull'intelligenza artificiale" consultabile al <https://eur-lex.europa.eu/legal-content/IT/TXT/?uri=CELEX%3A52018DC0795>.

⁷"A Definition of Artificial Intelligence: Main Capabilities and Scientific Disciplines." Shaping Europe's Digital Future, <https://digital-strategy.ec.europa.eu/en/library/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines>.

⁸ N. Musacchio, G. Guaita, A. Ozzello, M.A. Pellegrini, P. Ponzani, R. Zilich, A. De Micheli "Intelligenza Artificiale e Big Data in ambito medico: prospettive, opportunità, criticità." in The Journal of AMD, 2018, Vol. 21-3, pag. 204

Il termine algoritmo viene spesso utilizzato per identificare applicazioni di intelligenza artificiale, ma in realtà è corretto sottolineare che un sistema di I.A. è costituito dall'insieme di numerosi algoritmi che svolgono le varie funzioni per cui sono proposti.

E' evidente che i sistemi di I.A. siano complessi, stante la molteplicità di metodi e tecniche utilizzate ma anche la diversità dei gradi di autonomia delle macchine, la capacità di apprendimento, l'adattamento e la complessità intrinseca degli algoritmi stessi.

Per capire a fondo gli algoritmi è necessario partire dalla definizione offerta da Al-Khwarizimi⁹ che li definisce come "la risoluzione di un problema mediante la riduzione, semplificazione o riequilibrio di un processo".

L'algoritmo è un procedimento sistematico in cui vengono descritte operazioni di calcolo volte a semplificare un processo complesso.

Scopo principale dell'algoritmo è quello di ridurre i margini di errori che l'uomo potrebbe commettere impiegando il medesimo tempo.

Seguendo l'utile spiegazione di Burrell, un algoritmo consiste in due algoritmi distinti che operano in parallelo, uno "studente" e un "classificatore". Lo "studente" apprende da insiemi etichettati "dati di allenamento" che dettano la cosiddetta "verità di base".

I modelli derivati dai dati di allenamento sono continuamente ottimizzati dall'algoritmo durante il suo "apprendimento" fino a quando i risultati non diventano stabili.

Il "modello" finale viene quindi utilizzato dal classificatore anche per analizzare i nuovi dati di input denominati "caratteristiche" (ad es. pixel di nuovi schermi prostatici da diagnosticare) e determinare la classificazione dell'output come obiettivo effettivo.

L'output di solito include un punteggio di probabilità che riflette la fiducia dell'algoritmo rispetto a questa "previsione".

Il cosiddetto deep learning fa uso di reti neurali artificiali, una raccolta di semplici unità matematiche addestrabili (neuroni), che collaborano per calcolare una funzione complicata. I neuroni sono organizzati in livelli, in cui ogni livello apprende dal livello sottostante.

Quindi, salendo la scala, la rete si articola di modelli che aumentano continuamente la "densità semantica" nei dati.

Questo modo probabilistico di trasformare i dati in "conoscenza" che ha come base una "previsione", è ciò che viene definito "processo decisionale".

⁹ Matematico, astronomo, fisico che visse a cavallo dell' 800 d.C. e dedicò tutti i suoi trattati all'algebra e astronomia.

I metodi principali per addestrare l'agente intelligente sono il cosiddetto "machine learning" ossia il sistema di apprendimento automatico e il deep learning o apprendimento profondo. La differenza principale tra i due metodi si sostanzia nel processo di apprendimento dell'algoritmo.

Nel primo caso affinché l'agente intelligente possa dare una risposta o prendere una decisione, lo stesso si affida agli algoritmi di apprendimento che si basano su metodi statistici per ricondurre le esperienze in determinate categorie di comportamento umano e destinate a perfezionarsi nel tempo mediante l'acquisizione di dati sempre più aggiornati.

Il sistema di apprendimento automatico può essere definito come è un sistema di apprendimento basato sull'input iniziale di una grande quantità di dati.

L'algoritmo classificherà poi i dati inseriti grazie al riconoscimento di pattern preordinati come, ad esempio, negli aeroporti, uno scanner a raggi X individua gli oggetti pericolosi.

Nell'apprendimento automatico l'accuratezza dell'identificazione di determinate immagini e suoni dipenderà dalla quantità di dati e dalla capacità di elaborazione dell'apparecchiatura.

Ciò comporta, in particolare nel settore sanitario, problemi riguardanti il campo di riservatezza dei dati elaborati dall'algoritmo. Questo problema verrà affrontato nei prossimi paragrafi del presente lavoro di tesi.

I principali sistemi di apprendimento automatico consentono il riconoscimento mediante modelli ideali prestabiliti e si dividono in tre gruppi: ragionamento deduttivo, ragionamento induttivo e ragionamento abduttivo.

Il ragionamento deduttivo rappresenta il fondamento della logica matematica ed infatti viene utilizzato per dimostrare i teoremi matematici. Si parte da assunzioni vere e si inferiscono nuovi teoremi mediante dimostrazione. Secondo il ragionamento deduttivo la conclusione deve necessariamente essere conseguenza principale della premessa.

Il limite di questo tipo di ragionamento è che non permette di scoprire o prevedere fatti nuovi e non calcolabili.

Il ragionamento induttivo, invece, è una delle forme più semplici di machine learning supervisionato. Si basa esclusivamente sull'osservazione. Dato un primo insieme di esempi input-output, l'agente elabora ipotesi per ricostruire la funzione di causa ed effetto ignoti.

A differenza del ragionamento deduttivo, quello induttivo è di tipo probabilistico, infatti le conclusioni dipendono dal grado di probabilità delle informazioni contenute nelle premesse.¹⁰

¹⁰ Pasceri, G. *"Intelligenza Artificiale, Algoritmo E Machine Learning: La Responsabilità Del Medico e Dell'amministrazione Sanitaria."* Giuffrè Francis Lefebvre, 2021

Infine per quanto riguarda il ragionamento abduittivo, analogamente a quello induttivo è un ragionamento probabilistico però invece di generalizzare il risultato, ipotizza che alla conclusione a cui si giunge debba essere validata per via empirica.

Il termine Deep Learning si riferisce a un dominio di machine learning che utilizza reti neurali artificiali (Artificial Neuronal Networks - ANNs) analogamente al modo in cui il sistema nervoso umano elabora le informazioni. In altre parole, una ANN è composta da più nodi interconnessi e ciascun nodo esegue una serie di calcoli non lineari derivati da segnali di ingresso e nodi ad esso collegati.¹¹

Questi sistemi hanno raggiunto i risultati migliori in compiti che prima del Deep Learning potevano essere affrontati solo dagli esseri umani, come Computer Vision (essere in grado di identificare oggetti, modelli e altro nelle immagini) o Natural Language Processing o NLP (comprensione e creazione di testo e linguaggio naturale).

Tuttavia, il Deep Learning e le reti neurali artificiali hanno ancora molta strada da fare prima di poter raggiungere le capacità delle loro controparti umane.

Vediamo le principali sfide del Deep Learning:

I modelli di Deep Learning necessitano di enormi quantità di dati per raggiungere le loro prestazioni ottimali e per offrire l'eccellenza che ci aspetta da loro.

Tuttavia, avere così tanti dati non è sempre facile.

Una delle principali sfide del Deep Learning è essere in grado di fornire grandi prestazioni con molti meno dati di allenamento. I recenti progressi come il transfer learning o l'apprendimento semi-supervisionato stanno già facendo passi in questa direzione, ma non sono ancora sufficienti.

I dati sono dinamici, cambiano attraverso diversi fattori come l'ora, la posizione e molte altre condizioni.

Tuttavia, i modelli di Machine Learning, inclusi quelli di Deep Learning, sono costruiti utilizzando un insieme definito di dati (il set di addestramento) e funzionano bene fintanto che i dati che vengono successivamente utilizzati per fare previsioni.

Un'altra sfida del Deep Learning in futuro sarà superare questo problema e continuare a funzionare abbastanza bene quando gli verranno forniti dati che non corrispondono esattamente ai dati di addestramento, incorporando una sorta di conoscenza basata su regole, in modo che le procedure

¹¹ Mauloni S., Mazzanti M., Buscemi L. *“Focus: nuove tecnologie e risvolti medico legali - la nuova era dell'intelligenza artificiale: profili medico-legali - the new era of artificial intelligence: medico-legal aspects”* Rivista Italiana di Medicina Legale (e del Diritto in campo sanitario), fasc.3, 1 giugno 2022, pag. 679

logiche possano essere implementate e il ragionamento sequenziale utilizzato per formalizzare la conoscenza.

Il successo del Deep Learning deriva dalla possibilità di incorporare molti livelli nei nostri modelli, consentendo loro di provare un numero folle di combinazioni di parametri lineari e non lineari. Tuttavia, con più livelli aumenta la complessità del modello e abbiamo bisogno di più dati affinché questo modello funzioni correttamente.

Quando la quantità di dati che abbiamo è effettivamente inferiore alla complessità della rete neurale, allora si deve ricorrere a un approccio diverso come il suddetto Transfer Learning.

Inoltre, i modelli di Deep Learning troppo grandi, oltre a richiedere enormi quantità di dati su cui essere addestrati, utilizzano molte risorse di calcolo e possono richiedere molto tempo per l'addestramento.

I progressi sul campo dovrebbero anche essere orientati a rendere il processo di formazione più efficiente ed economico.

Ora che sono stati sottolineati i problemi principali di DL, è opportuno verificare i progressi più recenti e dove si sta dirigendo.

Utilizzati principalmente nel regno dell'elaborazione del linguaggio naturale (NLP), i meccanismi di attenzione e le architetture dei trasformatori hanno rivoluzionato il modo in cui le macchine capiscono e affrontano il testo e il linguaggio.

In parole povere, i meccanismi di attenzione fanno ciò a cui si riferisce il loro stesso nome, ossia consentono ai modelli di Deep Learning di prestare attenzione a una parte specifica del testo su cui stanno analizzando/elaborando/apprendo. Ciò significa che per un testo o una frase specifici, la rete neurale sceglie dinamicamente a quali parti del testo dare più peso e quali parti ricordare/memorizzare per un uso successivo.

Tra i sistemi di apprendimento dell'algoritmo è opportuno menzionare "l'apprendimento automatico senza supervisione". Tale tecnica si caratterizza per il fatto che l'algoritmo fornisce all'agente intelligente una serie di dati input che verranno organizzati e classificati in base ad elementi comuni.¹² Questo tipo di sistema deve affrontare problemi come il *clustering*¹³, in cui l'utilizzo di un pre-addestramento non supervisionato può consentire il raggruppamento di dati con caratteristiche simili, utilizzando la tecnica dell'associazione.

¹² Pasceri, G. "Intelligenza Artificiale, Algoritmo E Machine Learning: La Responsabilità Del Medico e Dell'amministrazione Sanitaria." Giuffrè Francis Lefebvre, 2021.

¹³ Un cluster è un insieme di oggetti che presentano tra loro delle similarità, ma che, per contro, presentano dissimilarità con oggetti in altri cluster.

Questo tipo di tecnica non appare idonea per alcuni campi, come ad esempio quello della sanità, in quanto procedendo per analogie potrebbero determinarsi, si pensi alla diagnostica per immagini, a falsi positivi o falsi negativi.

Questo metodo è simile a ciò che viene fatto nell'apprendimento del trasferimento, transfer learning, che viene utilizzato per riaddestrare un modello di rete neurale, allo scopo di risolvere un problema simile a quello per cui è stato progettato, mediante l'utilizzo di uno o più livelli dello stesso.

Il transfer learning, dunque, può essere molto utile nel caso in cui lo sviluppo del modello di rete neurale abbia portato a etichettare molti più dati rispetto al problema di iniziale interesse, e vi è una somiglianza nella struttura del nuovo problema da risolvere.

L'obiettivo, in parole semplici, è quello di sfruttare i dati della prima impostazione del modello neurale per estrarre informazioni che possono essere utili durante la fase di apprendimento del modello o per effettuare nuove previsioni in una seconda impostazione dello stesso.

L'apprendimento autocontrollato o "supervisionato" è l'insieme di algoritmi automatici più conosciuto e più utilizzato perché molto simile all'apprendimento di un bambino che apprende l'esempio, sperimentando e acquisendo conoscenza.

L'apprendimento supervisionato è una modalità di apprendimento in cui i dati sono etichettati e per ogni input, il rispettivo output è noto e utilizzato per insegnare all'algoritmo le regole del modello.

Ad esempio, supponiamo che si voglia creare un modello per riconoscere i semafori, il set di dati dovrebbe essere composto da luci verdi, gialle e rosse, ogni luce dovrebbe essere etichettata con il colore corretto.¹⁴

Già in questa fase potrebbe sorgere un dubbio, se avesse un altro colore, come reagirebbe il modello formato solo dai tre colori?

Un aspetto chiave dell'apprendimento supervisionato è che per produrre correttamente un dato un insieme di input, devono aver visto una situazione simile o prima.

Proprio come un bambino che non conosce il colore sarebbe indeciso. Tuttavia, l'algoritmo deve comunque dare una risposta per la quale, se errata, risponderà con uno dei tre colori che ha appreso e darà probabilità molto vicine associate a quelle possibili.

Tutto questo potrebbe essere paragonato al nostro tirare a indovinare a caso quando non conosciamo la risposta a un quesito.

¹⁴ Pasceri, G. *“Intelligenza Artificiale, Algoritmo E Machine Learning: La Responsabilità Del Medico e Dell'amministrazione Sanitaria.”* Giuffrè Francis Lefebvre, 2021

4. I Big Data e la medicina

I Big Data sono un “fenomeno culturale, tecnologico e scientifico” centrato sull’applicazione dell’apprendimento automatico a database di grandissime dimensioni.

Nel 1997, due ricercatori della NASA, Michael Cox David Ellsworth, usarono per primi "big data" per descrivere la loro difficoltà nell'affrontare la crescente quantità di dati generati nello studio dell'aerodinamica di aeromobili.¹⁵

Successivamente questo termine ha avuto un duplice significato: da un lato è stata data una definizione relativa a una grande mole di dati, dall'altro il suo significato è stato esteso anche all’insieme degli algoritmi utile per la sua elaborazione.

Il mondo dei Big Data ruota attorno alle famose “7 V”: Volume (grandi quantità di registri archiviati), Velocità (rapidità nella creazione e nel trasferimento dei dati), Varietà (dati strutturati e non strutturati), Veridicità (l'affidabilità di dati in termini di accuratezza), Valore (dati a più valori), Variabilità (dati il cui significato è in continua evoluzione), Varietà (diversi tipi di dati) e Visualizzazione (dati che siano accessibili e leggibili).¹⁶

Le organizzazioni sanitarie hanno oggi database di ingenti dimensioni, contenenti dati clinici, biologici, epidemiologici ed amministrativi; altri dati di possibile interesse medico possono provenire da fonti diverse, quali ad esempio i dispositivi indossabili che tracciano la mobilità, misurano variabili biologiche, monitorizzano gli stili di vita oppure le informazioni sui comportamenti individuali o richieste sulla salute dei social networks o dei motori di ricerca.

L'ascesa delle tecnologie sanitarie digitali ha portato a grandi quantità di "big data" qualitativi e quantitativi i vantaggi degli approcci "big data" vanno dai progressi nella biologia di base alla diagnosi assistita da computer.

Attraverso l’Intelligenza Artificiale si possono quindi analizzare ed interpretare i dati disponibili per formulare ipotesi ed ottenere risposte utili per una diagnosi precoce ed un trattamento migliore, in più sartorializzati sulla peculiarità del singolo paziente; in altri termini è possibile perseguire gli obiettivi

¹⁵ Mauloni S., Mazzanti M., Buscemi L., “Nuove tecnologie e risvolti medico legali - la nuova era dell'intelligenza artificiale: profili medico-legali - the new era of artificial intelligence: medico-legal aspects” Rivista Italiana di Medicina Legale (e del Diritto in campo sanitario), fasc.3, 1 giugno 2022, pag. 679.

¹⁶ Mauloni S., Mazzanti M., Buscemi L., “Nuove tecnologie e risvolti medico legali - la nuova era dell'intelligenza artificiale: profili medico-legali - the new era of artificial intelligence: medico-legal aspects” Rivista Italiana di Medicina Legale (e del Diritto in campo sanitario), fasc.3, 1 giugno 2022, pag. 679.

di efficacia, efficienza ed appropriatezza tipici della medicina attuale, attraverso l'utilizzo di tecnologie capaci di affinare ed innovare le classiche metodologie epidemiologiche su quantità di dati molto vaste.

L'analisi può riguardare sia i processi collegati ad esiti solidi, quali mortalità e morbilità, sia la soddisfazione dell'utente.

Altre applicazioni riguardano l'automatizzazione della diagnostica, la gestione economica e la ricerca.

Questa rivoluzione tecnologica è alla base della realizzazione della medicina delle 4P¹⁷ volta a superare il concetto della "cura delle malattie" per rivolgersi alla "gestione complessiva del benessere degli individui".

Il termine 4P Medicine è stato coniato dallo scienziato Leroy Hood.¹⁸ Le 4P stanno per prevenzione, predizione, personalizzazione e partecipazione, una rivoluzionaria architettura clinica basata su una visione generale e d'insieme del paziente.

Esiste una condivisa fiducia che l'utilizzo dei Big Data da fonti diverse porterà vantaggi per la cura dei pazienti, per la ricerca di terapie innovative, per l'attività degli operatori sanitari e per i gestori della salute pubblica.

Vi sono tuttavia alcuni aspetti che presentano delle criticità:

- La qualità dei dati, spesso provenienti da database nati per altri fini e non per la cura e la ricerca scientifica.
- La validità ed il corretto utilizzo dei dati.
- La disponibilità di dati clinici informatizzati ancora largamente incompleta (nel Regno Unito si calcola che gli ospedali saranno totalmente "senza carta" non prima del 2027).
- Il rispetto delle normative sulla privacy nella condivisione e trasmissione dei dati dei pazienti.

¹⁷ "Medicina Preventiva-Predittiva-Personalizzata-Partecipativa". È un termine coniato dallo scienziato Leroy Hood, convinto assertore del passaggio culturale dalla "cura delle malattie" alla "gestione complessiva del benessere degli individui". Hood sostiene che il concetto di paziente standard fra qualche anno verrà sostituito da quello di "individuo geneticamente unico", su cui cucire su misura l'approccio più corretto per preservare la sua salute o, meglio, il suo benessere. Questo nuovo, rivoluzionario concetto di medicina, è favorito da 3 tendenze: lo sviluppo della medicina dei sistemi, la disponibilità di dati generati dagli individui (IOT, social network) e la rivoluzione digitale.

¹⁸ Biologo americano che ha servito nelle facoltà del California Institute of Technology e dell'Università di Washington. Hood ha sviluppato strumenti scientifici rivoluzionari che hanno reso possibili importanti progressi nelle scienze biologiche e nelle scienze mediche.

- La difficoltà tecnica ed anche formale/ autorizzativa (proprietà del dato vs. trasparenza) di raccogliere dati da fonti diverse.
- La gestione delle grandi modificazioni nella pratica clinica e nella formazione professionale dei medici e degli altri operatori sanitari indotte dall'utilizzo dell'Intelligenza Artificiale, ad esempio per diagnosi e screening automatizzati.
- I problemi etici legati all'utilizzo dell'Intelligenza Artificiale.
- Il rapporto fra l'utilizzo dell'Intelligenza Artificiale ed il consolidato approccio della Evidence Based Medicine (EBM)¹⁹ nella Medicina Clinica.²⁰

La ricerca scientifica tradizionale si è sempre basata su raccolti rigorosamente su piccoli campioni, mentre i dati forniscono, a costi inferiori e facilmente accessibili, meno puliti ma sull'insieme della popolazione.

Da questa prospettiva, quindi, il volume dei dati potrebbe superare il suo disordine”, anche se i ricercatori potrebbero trovare eretico questo approccio epistemologico.

In altre parole, poiché la popolazione dello studio si avvicina al 100% anziché essere un 1% altamente e attentamente raccolto, il potere predittivo potrebbe essere maggiore.

I Big Data possono essere analizzati con tecniche artificiali, in grado di generare ipotesi dai dati stessi e dalle loro associazioni, per testare altre analisi in modo che possano confermare il loro possibile valore. Tuttavia, sottolineiamo che l'associazione statistica non significa causalità.

Restano comunque alcune perplessità:

- I dati devono essere realmente rappresentativi di una popolazione e sottopopolazione, come ad esempio soggetti seguiti da una particolare organizzazione, in cui, peraltro, la raccolta dei dati non può essere omogenea nei diversi contesti assistenziali; la qualità della imputazione e della codificazione deve essere garantita; i criteri diagnostici delle singole condizioni devono essere precisi ed inequivocabili ed anche omogenei nei diversi studi.

¹⁹ La Medicina fondata sulle prove, secondo la definizione di David L. Sackett, con Archibald Cochrane fra i “padri” della EBM, è “l’ integrazione delle migliori prove di efficacia clinica con la esperienza e l’abilità del medico ed i valori del Paziente”. Possiamo esprimere lo stesso concetto in un altro modo: la EBM è “l’uso cosciente, esplicito e giudizioso delle migliori evidenze (cioè prove di efficacia) biomediche al momento disponibili, al fine di prendere le decisioni per l’assistenza del singolo Paziente.

²⁰N. Musacchio, G. Guaita, A. Ozzello, M.A. Pellegrini, P. Ponzani, R. Zilich, A. De Micheli “*Intelligenza Artificiale e Big Data in ambito medico: prospettive, opportunità, criticità.*” in The Journal of AMD, 2018, Vol. 21-3, pag. 204

- I database amministrativi forniscono dati raccolti per scopi non clinici ed i dati delle cartelle elettroniche non sempre sono strutturati per una raccolta adeguata ai fini della ricerca scientifica e talora non sono raccolti con completezza.
- Gli algoritmi non possono riconoscere se modelli o associazioni trovate in assenza di un costruito teorico di base sono vere, spurie o falsate da errori sistematici. A differenza della EBM, il Machine Learning non ha un sistema per valutare la possibilità di errore sistematico o la qualità delle prove. Inoltre, non può distinguere variazioni di procedura giustificate o ingiustificate o analizzare le sequenze di decisioni cliniche.²¹

Esistono fondamentalmente due ordini di problemi.

I sistemi di supporto decisionale fondati sull'apprendimento automatico associano dati empirici ad una interpretazione categorica.

Potenziati conseguenze non intenzionali di questo approccio possono derivare dalla formalizzazione in un modello di decisione della mappatura tra i segni fisici che un medico può valutare e la loro “corretta” classe, identificata dagli osservatori. Nella pratica medica spesso gli osservatori non concordano sui segni diagnostici e sulla valutazione dei risultati e ciò non è solo soggettivo, ma legato ad una intrinseca ambiguità nei fenomeni clinici osservati. Quindi i dati di input per l'apprendimento artificiale non sono necessariamente univoci e precisi e ciò influenza inevitabilmente i risultati prodotti.²²

Ulteriore ricerca deve essere finalizzata a sviluppare e convalidare degli algoritmi di apprendimento automatico adatti a dati in entrata non precisi ed univoci come le informazioni mediche, anziché imporre un'idea di accuratezza e completezza dei dati non presente nei registri medici, in cui la qualità dei dati non è ottimale.

Affidarsi a sistemi di supporto decisionale fondati sull'apprendimento automatico potrebbe voler dire accettare che un dato digitale possa essere la rappresentazione completa ed affidabile di un fenomeno clinico; ciò diviene problematico se i medici perdono la coscienza della esistenza di elementi clinici non inclusi in una registrazione digitale. Ciò può portare ad una cattiva interpretazione delle risposte provenienti da un sistema automatico ed alla perdita di un approccio

²¹ N. Musacchio, G. Guaita, A. Ozzello, M.A. Pellegrini, P. Ponzani, R. Zilich, A. De Micheli “*Intelligenza Artificiale e Big Data in ambito medico: prospettive, opportunità, criticità.*” in The Journal of AMD, 2018, Vol. 21-3, pag. 204

²² . Musacchio, G. Guaita, A. Ozzello, M.A. Pellegrini, P. Ponzani, R. Zilich, A. De Micheli “*Intelligenza Artificiale e Big Data in ambito medico: prospettive, opportunità, criticità.*” in The Journal of AMD, 2018, Vol. 21-3, pag. 204

olistico al paziente nei suoi aspetti psicologici, relazionali e sociali ed anche ad una decontestualizzazione dagli aspetti organizzativi nella interpretazione dei dati.

Ad esempio, un modello prognostico realizzato con l'utilizzo dell'Intelligenza Artificiale ha dato l'indicazione che i malati asmatici ricoverati per polmonite muoiono meno dei non asmatici²³, dato chiaramente poco accettabile per i clinici. Invero, il modello utilizzato non aveva considerato che i malati con asma sono direttamente ricoverati in terapia intensiva per prevenire complicazioni e questa modalità organizzativa non poteva essere inserita nel modello prognostico.

In queste prospettive un'agenzia indipendente che certifichi i modelli di previsione prima di diffonderli nella pratica clinica potrebbe assicurare che le analisi predittive mantengano le loro promesse di migliori risultati per singoli pazienti e l'intero sistema sanitario.

I risvolti giuridici del rapporto tra medico e macchina verranno meglio analizzati nel capitolo 3 del presente elaborato.

5. Il fenomeno dei black box applicato al settore sanitario

I recenti progressi nello sviluppo di algoritmi di apprendimento automatico (ML), combinati con l'enorme quantità di dati utilizzati per addestrarli, hanno cambiato radicalmente la loro utilità e la portata delle applicazioni. Gli strumenti software basati su questi algoritmi sono ora abitualmente utilizzati nei sistemi di giustizia penale, nei servizi finanziari, nella ricerca e persino nelle attività aziendale.

Molte decisioni che riguardano importanti aspetti della nostra vita sono ora prese da algoritmi piuttosto che da esseri umani. Chiaramente, ci sono molti vantaggi in questa trasformazione. Le decisioni umane sono spesso di parte e talvolta semplicemente errate. Gli algoritmi sono più economici e più facili da adattare alle mutevoli circostanze.

Diverse tecniche possono essere utilizzate per generare modelli predittivi.²⁴ Una distinzione importante riguarda gli algoritmi che vengono definiti "black box" o scatole nere, e gli algoritmi AI che sono "trasparenti". La prima parte di entrambi i tipi di analisi è la stessa: generare un modello

²³ Musacchio, G. Guaita, A. Ozzello, M.A. Pellegrini, P. Ponzani, R. Zilich, A. De Micheli "Intelligenza Artificiale e Big Data in ambito medico: prospettive, opportunità, criticità." in The Journal of AMD, 2018, Vol. 21-3, pag. 204.

predittivo che consente di inserire informazioni su nuovi individui per vedere se l'individuo esaminato ha maggiori probabilità di cadere in errore in base a ciò che il modello ha appreso dai dati precedenti. Il modello fornisce una risposta sì/no (ad esempio, se c'è un rischio).

Un modello predittivo di tipo "scatola nera" ci dice se un soggetto è a rischio, ma non ne spiega il motivo. Questo è tipico dei modelli generati utilizzando tecniche come le reti neurali o le support vector machine (SVM): sono delle "scatole chiuse" che, in virtù di complesse funzioni matematiche, ci danno una risposta di tipo sì/no senza aggiungere ulteriori informazioni. Al contrario, le tecniche di intelligenza artificiale trasparente, oltre alle risposte predittive sì/no, espongono i parametri su cui si basano le previsioni, cioè spiegano alle persone che utilizzano il modello qual è la logica interna utilizzata dal modello stesso per fornire la sua valutazione.

Questo è tipico di algoritmi specifici chiamati "Logical Learning Machines", che, inoltre, descrivono in linguaggio naturale le regole con cui il modello predittivo "ragiona".²⁵

Le operazioni "scatola nera" possono essere ricondotte al concetto di "processo decisionale automatizzato": software che prende decisioni in modo autonomo, senza fornire spiegazioni.

Di contro, le funzionalità trasparenti consentono il cosiddetto "processo decisionale aumentato": invece di prendere una decisione in modo autonomo, il software consente a un esperto umano in un campo, come un medico, di prendere una decisione sulla base dell'evidenza delle correlazioni trovate nel modello e utilizzare queste informazioni aggiuntive combina questo con il bagaglio di conoscenze precedenti.

L'utilizzo di strumenti trasparenti e la conoscenza dei parametri alla base dei modelli predittivi consente di estrarre il massimo valore dall'intelligenza artificiale. In altre parole, la conoscenza generata dal modello consente di correggere la situazione e di evitare che vada verso un esito indesiderato. Ad esempio, se i fattori di rischio individuati dal modello sono fattori suscettibili di intervento con farmaci o diversi stili di vita, è possibile alterare l'evoluzione della patologia di un determinato individuo, eventualmente determinando una situazione dalle conseguenze meno nefaste. E' comunque corretto sottolineare che non sono mancati diversi casi di pregiudizio e discriminazione dettati nel processo decisionale algoritmico, particolarmente inquietanti ad esempio nel caso della giustizia criminale. Naturalmente, una volta scoperti, tali pregiudizi possono essere rimossi e gli algoritmi possono essere convalidati come non discriminatori prima di essere distribuiti. Ma c'è ancora un diffusa perplessità, in particolare tra gli esperti di diritto, sull'uso di questi algoritmi.

²⁵ Musacchio N, Baccetti F, Giorda CB, Guaita G, Morviducci L, Nreu B, Ozzello A, Ponzani P, Rossi A, Zilich R. *"Intelligenza artificiale: la posizione e le strategie di AMD."* in The Journal of AMD, 2022, Vol. 25, pag. 219-226.

La maggior parte di questi algoritmi sono ad autoapprendimento e i loro progettisti hanno poco controllo sui modelli generati dai dati di training.

In effetti, gli informatici in passato non erano molto interessati a studiare questi modelli perché erano (e sono) spesso straordinariamente complessi (il motivo per cui vengono spesso indicati come "scatole nere"). L'approccio standard era che finché un algoritmo funzionasse correttamente, nessuno si preoccupava di analizzarne il funzionamento.²⁶

L'approccio è cambiato una volta che gli strumenti basati su algoritmi ML sono diventati onnipresenti e hanno iniziato a influenzare direttamente la vita della gente comune.

Se la decisione su quanti anni si trascorrerà in prigione è presa da un algoritmo, il condannato dovrebbe avere il diritto di sapere come viene presa questa decisione. In altre parole, vi è una chiara necessità di trasparenza e responsabilità degli algoritmi di processo decisionale automatico (ADM). Negli ultimi anni, molti articoli pubblicati hanno affrontato l'interpretabilità (variamente definita) dei modelli generati dagli algoritmi ML. È stato sostenuto che l'interpretabilità non è una nozione monolitica. Di conseguenza, la soggettività di ogni interpretazione, dovuta ai diversi livelli di comprensione umana, implica che ci deve essere una moltitudine di dimensioni che insieme costituiscono l'interpretabilità.

Diventa evidente che i modelli creati dagli algoritmi ML sono intrinsecamente così complessi che non possono essere "spiegati" in modo significativo a un utente ordinario di tali sistemi. Invece, piuttosto che guardare "dentro" un algoritmo, si propone di mettere a fuoco sulla sua correttezza statistica. Un modo promettente per raggiungere questo obiettivo potrebbe essere quello di introdurre meccanismi di registrazione degli eventi e schemi di certificazione, che sono attualmente utilizzati con grande successo nel settore IT.

Come è stato già rappresentato, l'interpretabilità dell'algoritmo è un compito impegnativo per i loro progettisti. Di solito si distinguono tre ostacoli alla trasparenza degli algoritmi in generale:

- 1) occultamento intenzionale il cui obiettivo è la protezione della proprietà intellettuale;
- 2) mancanza di alfabetizzazione tecnica da parte degli utenti;
- 3) opacità intrinseca che deriva dalla natura dei metodi ML.

La soluzione più ovvia alla seconda barriera sarebbe l'attuazione di programmi educativi volti a trasferire le conoscenze sul funzionamento delle moderne tecnologie. Ciò potrebbe essere ottenuto con programmi di educazione al pensiero computazionale più incisivi e con la messa a disposizione di esperti indipendenti che forniscano consulenza a coloro che sono interessati dal processo decisionale algoritmico. L'efficacia di questa soluzione, tuttavia, è discutibile: anche se fosse

possibile migliorare l'educazione all'alfabetizzazione tecnica (cosa che sembra molto improbabile istesse precedenti esperienze in questo campo), rimane comunque il fatto che l'80% della popolazione ha completato la propria istruzione molti anni fa.²⁷

Come soluzione all'ultima barriera, ovvero la mancanza di trasparenza relativa alla natura dei metodi di ML, dovrebbe essere sufficiente una sorta di raccolta di prove basata sulla registrazione dei parametri chiave dell'algoritmo.

In effetti, la raccolta di questo tipo di dati aiuterebbe certamente a capire come un sistema sia arrivato a una specifica decisione. Detto questo, sarebbe comunque del tutto irrealistico aspettarsi che un profano afferri questi concetti.

Negli ultimi anni si è lavorato molto sulla spiegazione dei modelli "a scatola nera". Alcuni di questi lavori sono stati rivolti specificamente agli esperti. L'interpretabilità di un modello è un elemento chiave di una procedura di validazione robusta in applicazioni come la medicina o le auto a guida autonoma.

Spiegare il modello della scatola nera di un algoritmo è molto più difficile di quanto si creda.

Per illustrare meglio questo caso, si riporta un recente lavoro sulla progettazione di un sistema di raccomandazione di canzoni per KKBOX, il principale fornitore di servizi di streaming musicale in Asia.²⁸

KKBOX ha fornito un set di dati di addestramento che consisteva in informazioni provenienti da sessioni di ascolto per ogni coppia utente-canzone all'interno di un periodo di tempo specifico.

Le informazioni a disposizione dell'algoritmo comprendono informazioni sugli utenti, come il numero di identificazione, l'età, il sesso, ecc. e sulle canzoni, come la lunghezza, il genere, il cantante, ecc. I dati di addestramento e di test sono stati selezionati dalla cronologia degli ascolti degli utenti in un determinato periodo di tempo e contano rispettivamente circa 7 e 2,5 milioni di coppie uniche utente-canzone.

La qualità delle previsioni di un sistema di raccomandazione si basa su due fattori principali: le caratteristiche predittive disponibili dai dati passati (ad esempio, le canzoni più ascoltate dall'utente) e un algoritmo di apprendimento efficace.

²⁷ Gryz, Jarek, Marcin Rojszczak. "Black box algorithms and the rights of individuals: No easy solution to the "explainability" problem." *Internet Policy Review*, 2021 vol. 10, n. 2, 2021, pag. 1-24 <https://doi.org/10.14763/2021.2.1564>.

²⁸ Gryz, Jarek, Marcin Rojszczak. "Black box algorithms and the rights of individuals: No easy solution to the "explainability" problem." *Internet Policy Review*, 2021 vol. 10, n. 2, 2021, pag. 1-24 <https://doi.org/10.14763/2021.2.1564>.

Molto spesso, queste caratteristiche sono solo implicite nei dati di addestramento e l'algoritmo non è in grado di estrarle da solo.

L'ingegneria delle caratteristiche è un approccio che sfrutta la conoscenza del dominio di un esperto per estrarre dal set di dati le caratteristiche che dovrebbero generalizzare bene ai dati non visti del set di test. La qualità e la quantità delle caratteristiche hanno un impatto diretto sulla qualità complessiva del modello.

In questo caso, sono state create (o estratte, perché non erano esplicitamente presenti nei dati) alcune caratteristiche statistiche, tra cui il numero di sessioni per utente, il numero di canzoni per sessione e la durata della registrazione di un utente a KKBOX.

E questo è il punto chiave: alcune di queste caratteristiche si sono rivelate estremamente importanti per determinare i gusti dell'utente e, di conseguenza, le raccomandazioni fornite. Ma va sottolineato che nessuna di queste caratteristiche era esplicitamente presente nei dati originali. Il paradosso è che se qualcuno chiedesse di spiegare il funzionamento del modello, la risposta dovrebbe basarsi su caratteristiche non presenti nei dati di partenza.

Tuttavia accade che i modelli ADM sono spesso ancora più complessi del sistema descritto sopra. L'apprendimento automatico è guidato dall'euristica e nessuno si aspetta prove matematiche rigorose della correttezza dei suoi algoritmi.

Spesso accade che, se un modello generato da un algoritmo non classifica correttamente i dati di prova, il progettista vi sovrappone un altro strato algoritmico nella speranza che migliori i risultati.

A questo punto nessuno è in grado di spiegare il risultato. Come ha detto Ali Rahimi²⁹ in un recente intervento alla Conference on Neural Information Processing Systems: "l'apprendimento automatico è diventato un'alchimia, molti progettisti di reti neurali usano una tecnologia che non capiscono veramente". Se le persone che progettano questi algoritmi non li capiscono, come possono gli altri?³⁰

Nella prospettiva medica non è accettabile un modello di "black box" in cui il razionale dell'input generato sia imperscrutabile o comunque esplicitabile solo a fronte di costose e lunghe analisi matematiche. Data la diffusione di queste nuove metodiche di analisi, sarà sempre più importante poter osservare e raffrontare in modo efficiente i differenti modelli aperti in modo da identificare i fattori di confondimento e scegliere il modello migliore.

²⁹ Rahimi, A. (2017). *NIPS 2017 Test-of-Time Award presentation*. Conference on Neural Information Processing Systems. <https://www.youtube.com/watch?v=ORHFOnaEzPc>

³⁰ Gryz, Jarek, Marcin Rojszczak. "Black box algorithms and the rights of individuals: No easy solution to the "explainability" problem." *Internet Policy Review*, 2021 vol. 10, n. 2, 2021, pag. 1-24 <https://doi.org/10.14763/2021.2.1564>.

Alla luce di ciò, ci si deve muovere verso sistemi di apprendimento artificiale in grado di offrire automaticamente spiegazioni e fornire ai medici strumenti di visualizzazione interattivi per esplorare il significato delle varie variabili di esposizione. Questa esigenza è diffusa non solo in medicina, ma in tutti gli ambiti in cui si utilizza l'intelligenza artificiale. In effetti, l'orientamento tecnologico sta cercando di fornire modelli di previsione più "trasparenti".

Disporre di strumenti tecnologici molto potenti non esonera i medici dall'acquisire maggiori competenze nel quantificare e valutare il valore degli strumenti basati sull'intelligenza artificiale nella pratica clinica e nella ricerca. Da questo punto di vista dovrebbe iniziare una formazione accademica speciale.

E' quindi chiaro che l'apprendimento automatico non è una bacchetta magica che può trasformare i dati in oro, ma uno strumento prezioso, sempre più necessario per il sistema sanitario moderno. La complessità della medicina oggi supera la capacità della mente umana.

Gli algoritmi a scatola nera prendono decisioni che influenzano le vite umane.

Questa tendenza non è destinata a cambiare nei prossimi anni. Le decisioni automatiche saranno prese non solo su scala sempre maggiore, ma anche con un'intensità sempre maggiore; di conseguenza, aumenterà anche la pressione sull'opinione pubblica per sviluppare meccanismi di controllo efficaci, compresi quelli che consentono di mettere in discussione le decisioni prese nei singoli casi.

Una soluzione collaudata, utilizzata dai progettisti di sistemi informatici nei casi in cui era necessario tracciare (ricostruire) il funzionamento di un algoritmo in una fase successiva, è la registrazione dei parametri di elaborazione significativi. Un esempio tipico di tale meccanismo sono i registratori di volo, ossia gli elementi chiave utilizzati per determinare il corso degli eventi di volo.

La presente proposta mira, pertanto, a introdurre l'obbligo di registrare (log) le ragioni delle decisioni prese da un algoritmo di ML.

I sostenitori di tale soluzione sottolineano la capacità di tracciare il corretto funzionamento del sistema e quindi l'accuratezza delle conclusioni raggiunte, si tratta di una sorta di "trasparenza qualificata": fornire ai singoli, agli esperti e alle autorità di regolamentazione diverse, ma appropriate, serie di informazioni relative al processo decisionale algoritmico.

La registrazione dei parametri rilevanti è relativamente semplice da implementare, non aumenta i costi di implementazione e manutenzione del sistema e non richiede procedure di validazione lunghe. Si tratta di vantaggi importanti perché, quando si considerano le proposte relative al rispetto dei requisiti normativi, non bisogna perdere di vista le loro conseguenze economiche.

I sistemi di ML sono per lo più sviluppati per un'applicazione globale. L'introduzione di normative la cui attuazione richiederebbe costi significativi a carico dei fornitori di tecnologia potrebbe portare a

una distorsione della concorrenza di mercato o alla delocalizzazione dei fornitori in giurisdizioni in cui tali normative non sono state attuate.

Oltre a essere semplice da implementare, la registrazione dei parametri del sistema può essere facilmente protetta in modo crittografico per garantire la coerenza e l'integrità dei dati registrati. Tenendo conto del tipo di sistema di ML o della sensibilità dei dati trattati, i log possono essere mantenuti da un fornitore di servizi specifico o da una terza parte fidata, evitando il rischio che i dati vengano modificati senza autorizzazione.

Inoltre, non vi è alcun ostacolo all'archiviazione di tali dati in sistemi supervisionati da enti pubblici; in questo modo, i parametri rilevanti, ad esempio, di un sistema di credit scoring³¹ basato su macchine, potrebbero essere archiviati in modo sicuro sotto la supervisione di un'autorità di vigilanza dei mercati finanziari.

Questo, a sua volta, apre la possibilità di introdurre requisiti specifici per il settore che definiscano un insieme minimo di parametri che devono essere registrati dai sistemi decisionali automatici e utilizzati per la fornitura di servizi nei mercati regolamentati.

In base a questo approccio, una persona che contesti la correttezza di una decisione presa o che desideri esercitare il proprio diritto alla spiegazione di una decisione automatica (articolo 22 del GDPR) avrebbe accesso all'insieme dei parametri chiave che hanno influenzato la decisione finale. A sua volta, l'autorità di controllo potrebbe avere accesso a un insieme più ampio (e più dettagliato) di parametri con cui analizzare non solo i singoli casi, ma anche la regolarità e la legalità del funzionamento dell'intero sistema.

La soluzione delineata sopra ha i suoi punti deboli. Innanzitutto, non può essere applicata a tutti i tipi di algoritmi di apprendimento automatico, in particolare alle reti neurali profonde con pesi legati a caratteristiche e interazioni complesse che non sono direttamente interpretabili, e quindi non è possibile registrare argomenti interpretabili dall'utente.

Inoltre, i sistemi di ML non sono 'statici': con nuovi dati, il modello di previsione generato da un algoritmo cambia. Di conseguenza, il processo di inferenza verrà modificato (ad esempio, verranno inclusi nuovi parametri o omessi parametri preesistenti) e cambieranno anche i meccanismi di registrazione degli eventi.

Nelle soluzioni informatiche tradizionali, è l'utente principale del sistema che determina l'insieme dei dati da registrare e indica anche la frequenza con cui tale registrazione deve essere effettuata. Sia

³¹ Un sistema usato dalle banche e dagli intermediari finanziari per valutare la solvibilità del consumatore. Il sistema combina tra loro una serie di informazioni al fine di pervenire ad un punteggio di accettazione (da parte del soggetto finanziatore) circa il rischio di credito del richiedente in un determinato arco di tempo.

l'ambito dei dati che la frequenza di registrazione del ML sono criteri che non possono essere definiti in anticipo. In pratica, è il sistema stesso (o uno dei suoi componenti) che deve essere progettato per determinare quali parametri devono essere registrati e quando. Tuttavia, ciò va contro l'idea alla base di questa salvaguardia: garantire la trasparenza.

Poiché non è lo sviluppatore a stabilire criteri rigidi e immutabili per la registrazione dei parametri chiave, ma il sistema stesso, questo meccanismo potrebbe essere soggetto a errori o a manipolazioni esterne.

Di conseguenza, sarebbe necessaria una valutazione formale del processo di registrazione stesso. In altre parole, il tentativo di risolvere il problema della trasparenza di un sistema ML verrebbe sostituito dal problema di garantire la trasparenza del sottosistema di registrazione degli eventi.

Un altro limite di questa soluzione è il contesto di analisi, che è difficile da prendere in considerazione. Va ricordato che il funzionamento di un algoritmo dipende non solo dai dati di ingresso e dalle procedure interne di elaborazione (il cui risultato è anche facile da salvare), ma anche dalle analisi precedenti, cioè dall'intero albero delle decisioni prese in precedenza.

La comprensione del risultato attuale di un algoritmo può quindi richiedere la revisione di un'enorme base di conoscenza che descrive le decisioni precedenti prese dal sistema. Senza queste informazioni, il semplice salvataggio dei parametri correnti utilizzati nell'inferenza potrebbe non consentire di ricostruire (e quindi verificare la correttezza e l'equità) dell'inferenza eseguita.

Quanto più un algoritmo è basato su meccanismi di apprendimento automatico, tanto più questo problema renderà difficile l'uso del logging come strumento per garantire la trasparenza del sistema.

Un terzo limite da discutere è la relazione non ovvia tra i parametri memorizzati e la logica interna di un algoritmo. Anche ipotizzando che i due ostacoli precedenti possano essere superati e che la registrazione dei parametri chiave permetta di riprodurre in modo completo e preciso lo stato iniziale e i risultati delle successive fasi di elaborazione, il problema dell'accesso alla logica interna di un algoritmo diventerà poi evidente. I sistemi di ML, come altre tecnologie altamente specializzate, sono soggetti alla tutela della proprietà intellettuale.³²

L'efficacia della protezione dei vari componenti della tecnologia AI è un problema significativo che incide sulla crescita di questo mercato. Senza l'accesso al codice sorgente - e quindi alla logica di un

³² Gervais, D. “*Is Intellectual Property Law Ready for Artificial Intelligence?*” GRUR International, 2020 vol. 69, n. 2, pag. 117–118, <https://doi.org/10.1093/grurint/ikz025>.

algoritmo di IA - anche i parametri dettagliati del suo funzionamento non saranno sufficienti per comprendere appieno il processo decisionale di cui si vuole valutare la correttezza.

Un'altra questione da chiarire è l'adeguatezza di questa misura nel raggiungere lo scopo prefissato. Infatti, i sostenitori della trasparenza del trattamento si aspettano che venga garantita l'affidabilità (credibilità) del funzionamento degli algoritmi. Sembra tuttavia che garantire la trasparenza del sistema non sia sempre una garanzia sufficiente dell'affidabilità del trattamento e quindi della tutela dei diritti dell'individuo. La garanzia della correttezza del trattamento deve comprendere non solo la conferma della correttezza del trattamento effettuato, ma anche la sua conformità a standard legali o etici. Tenendo conto di queste ulteriori limitazioni, può risultare che un sistema informatico correttamente funzionante, che identifica relazioni oggettivamente corrette tra i dati, non possa essere considerato affidabile. Non sarà possibile rivelare questa limitazione solo registrando i parametri di elaborazione.

Questi parametri da soli non riveleranno un difetto relativo ai dati esterni su cui si basa un algoritmo. Non ci si deve quindi aspettare che venga sviluppato un unico schema di certificazione universalmente accettato per i sistemi ADM.

È anche improbabile che tale standard uniforme venga sviluppato all'interno dell'UE. Il motivo non è solo la mancanza di consenso tra gli Stati membri sulla necessità di stabilire una regolamentazione comunitaria in questo settore, ma anche la diversa maturità digitale dei singoli mercati nazionali. Si sta sviluppando infatti una serie di diverse tutele legali da applicare in particolari Paesi dell'UE per garantire che lo sviluppo dinamico della tecnologia - compresa la diffusione dell'ADM - non influisca negativamente sull'area dei diritti fondamentali.

Questa tendenza e il problema dell'attuazione del diritto alla spiegazione delle decisioni prese automaticamente sono le principali aree di attività legislativa che verrà esaminata nei paragrafi successivi.

6. Sistemi robotici e l'agente intelligente in sanità

L'Association for Computing Machinery (ACM) categorizza la tecnologia dell'Intelligenza Artificiale in tre classi:

1. tecniche di IA: modelli matematici avanzati quali, ad esempio, il machine learning, che consentono di eseguire compiti che usualmente sono svolti dall'uomo;
2. applicazioni funzionali di IA: la percezione visiva, elaborazione del linguaggio naturale ecc.;

3. campi di applicazione dell'IA: le diverse discipline dove le tecniche di IA o le applicazioni funzionali possono giocare un ruolo importante. In medicina, tali ambiti applicativi sono rappresentati da: cura del paziente, diagnostica, ricerca e gestione.³³

Il successo dell'uso di agenti intelligenti nel settore sanitario ha spinto i ricercatori ad applicare questo paradigma emergente dell'ingegneria del software in applicazioni più avanzate e complesse.

Il principale fattore di successo è la mappatura naturale dei problemi medici del mondo reale nel mondo cibernetico.

L'architettura multi-agente può facilmente modellare i sistemi sanitari eterogenei, distribuiti e autonomi.

I sistemi multi-agente sono stati applicati a singole attività sanitarie come un sistema medico basato sulla conoscenza a sistemi complessi basati su più componenti come un'unità sanitaria completa.

L'uso di sistemi multi-agente nel settore sanitario ha anche aperto la strada alla scoperta di nuove applicazioni come i sistemi sanitari personalizzati e socializzati.

Questo uso versatile di sistemi multi-agente ha anche posto nuovi problemi a ricercatori come sicurezza, comunicazione e diverse problematiche sociali.

Questo lavoro di tesi ha esaminato e confrontato una serie di modelli di IA applicati alla sanità per individuare i pregi e le carenze di tali sistemi. E' necessario discutere le future linee guida e le direzioni di ricerca per i professionisti del settore.

La salute umana, essendo capitale umano, è di primaria importanza e punto focale di tutti gli sforzi umani.

Pertanto, la maggior parte dell'attenzione e degli sforzi umani sono diretti all'assistenza sanitaria e al benessere sociale.

Il calcolo o l'informatizzazione sono tra le tecnologie che hanno dimostrato maggiore impatto sulla vita umana fornendo sistemi sanitari assistivi, collaborativi e automatizzati.

Lo sviluppo dei sistemi sanitari se da un lato ha migliorato considerevolmente l'attività medica dall'altro ha suscitato non poche problematiche relative alla privacy dei dati dei pazienti, le questioni culturali, legali e tecniche.

È necessario chiedersi se i sistemi sanitari convenzionali basati sui paradigmi informatici tradizionali siano all'altezza di modellare la natura complessa e distribuita dei sistemi sanitari.

³³Mauloni S., Mazzanti M., Buscemi L., *“Nuove tecnologie e risvolti medico legali - la nuova era dell'intelligenza artificiale: profili medico-legali - the new era of artificial intelligence: medico-legal aspects”* Rivista Italiana di Medicina Legale (e del Diritto in campo sanitario), 2022 fasc.3, pag. 679.

I sistemi basati su agenti sono uno dei nuovi paradigmi di analisi, progettazione, sviluppo e distribuzione per sistemi complessi e distribuiti.

In questo paradigma, i componenti software non sono strettamente accoppiati tra loro e possono funzionare in modo autonomo, anche in ambienti imprevedibili.

L'uso di agenti intelligenti nei sistemi sanitari ha fornito il valore aggiunto alla loro funzionalità primaria. In letteratura, l'uso di agenti software intelligenti è stato proposto per affrontare una varietà di problemi medici e sanitari, ad esempio la programmazione dei pazienti, la gestione dei trapianti di organi e tessuti, l'assistenza comunitaria, accesso alle informazioni sui pazienti e sul trattamento, supporto alle decisioni, formazione e gestione ospedaliera, assistenza agli anziani, auto-cura e monitoraggio automatico della salute³⁴.

È chiaro che l'uso di componenti software autonomi nei sistemi medici è diventato una delle aree di ricerca più interessanti.

Un software che funziona autonomamente in un dato ambiente, ottiene la percezione dall'ambiente, lo traduce in una forma rappresentabile basata sulla sua base di conoscenze ed esegue azioni per raggiungere i suoi obiettivi di progettazione, è noto come agente software³⁵.

In un sistema multi-agente, un determinato compito è suddiviso in sotto-compiti che sono ulteriormente assegnati a diversi agenti in base a un meccanismo di coordinamento supportato da protocolli di comunicazione e negoziazione.

Ciascun agente produce output in base all'attività assegnata che viene combinata con l'output di altri agenti per creare l'output di risultato di un'attività più ampia. In un sistema multi-agente, ogni agente può essere progettato in modo diverso. Questa eterogeneità crea flessibilità nel sistema per operare in un ambiente eterogeneo utilizzando metodi di coordinamento.

Non esiste un'unica tassonomia ben definita dei sistemi multi-agente.

Possono essere classificati in tante classi quante sono le caratteristiche richieste.

Considerando varie caratteristiche e aspetti, i ricercatori hanno classificato i MAS in diversi gruppi. I criteri di raggruppamento possono includere lo stile di condivisione delle risorse, la descrizione funzionale, la descrizione strutturale, la descrizione comportamentale e il meccanismo di coordinamento.

³⁴ Moreno A. Nealon J. L. *“Applications of Software Agent Technology in the Health Care Domain”* Birkhauser Verlag, 2003.

³⁵ Nwana, Hyacinth S. *“Software Agents: An Overview.”* The Knowledge Engineering Review, 1996 vol. 11, n. 3, pp. 205–244., <https://doi.org/10.1017/s026988890000789x>.

Stakeholder e componenti di un sistema sanitario possono essere considerati agenti.

Nel sistema sanitario, gli agenti software devono affrontare alcuni problemi tecnici aggiuntivi quali quelli che riguardano la sicurezza, l'affidabilità, la riservatezza e l'efficienza.

I criteri di raggruppamento per gli agenti intelligenti in sanità includono la gestione dei dati medici, la preelaborazione dei dati, la raccolta dei dati, la comunicazione e la sicurezza dei dati, l'intelligenza medica ambientale, la base di conoscenze e i sistemi di supporto alle decisioni, i sistemi operativi per l'assistenza sanitaria, la pianificazione e la gestione delle risorse sanitarie.

I sistemi sanitari multifunzione sono considerati una categoria separata che coinvolge più di un sottodominio per formare un unico sistema integrato.

È opportuno riportare sinteticamente alcuni sottodomini aggregati in sanità:

1. Pianificazione e gestione delle risorse: questi sistemi si concentrano sulla pianificazione dei processi medici, sul monitoraggio del personale e sulla misurazione delle prestazioni, sul monitoraggio della salute dei pazienti, sulla gestione delle risorse ospedaliere e cliniche.
2. Sistemi di supporto alle decisioni/Sistemi di base di conoscenza: questa categoria include tali sistemi che utilizzano la knowledge base (KB) e applicano alcuni tipi di tecniche di analisi dei dati (utilizzando l'IA o l'apprendimento automatico), algoritmi di riconoscimento dei modelli e potrebbero anche utilizzare spesso tecniche di inferenza della conoscenza.
3. Sistemi di gestione dei dati: questi sistemi si concentrano principalmente sull'estrazione, la rappresentazione, l'organizzazione, l'archiviazione, il recupero e la presentazione dei dati sanitari.
4. Sistemi di assistenza remota/autoassistenza: quest'area comprende sistemi progettati per il monitoraggio automatico del paziente in remoto e l'autoassistenza del paziente.
5. Sistemi multifunzione: la revisione della letteratura ha rilevato che vari sistemi svolgono più attività relative a una soluzione sanitaria completa. E questi sistemi possono essere composti da due o più sottodomini indicati sopra, quindi sono stati classificati sotto questo titolo.

Come già evidenziato, i sistemi di pianificazione e gestione delle risorse nel campo sanitario si concentrano principalmente sul monitoraggio e la programmazione dei dati, sulla pianificazione dei processi medici, sul monitoraggio e sulla gestione del personale sanitario, sulla misurazione delle prestazioni, sulla gestione delle risorse ospedaliere e cliniche.

Sono diversi i sistemi MAS principali utilizzati per una pianificazione ottimale delle risorse in base alle condizioni del paziente.

Si pensi alla modellazione basata su agenti e sulla simulazione del triage del pronto soccorso³⁶, ossia un sistema sviluppato per il triage dei pazienti che si presentano al pronto soccorso dell'ospedale e sono in coda per la fornitura di servizi medici. Il sistema propone la valutazione della coda di attesa per riorganizzarla in base alla gravità delle condizioni dei pazienti e per gestire l'allocatione delle risorse presenti (medico o infermiere) in modo ottimale, utilizzandole prima per i casi urgenti.

Questo sistema MAS si basa su 9 agenti.

L'agente simulatore di pazienti per la generazione dei dati si basa su quelli del caso di studio. Questi dati vengono forniti all'agente paziente che agisce come paziente nella vita reale. I dati vengono inoltrati agli agenti del gestore code di triage e re-triage. L'agente di gestione delle code di triage prepara la coda dei pazienti in base alle loro condizioni di deterioramento con l'aiuto di un dispositivo di accesso al triage che registra il livello di gravità di un paziente in base allo stato delle condizioni cliniche dell'agente.

L'agente di gestione delle code di ri-triage gestisce il livello di gravità di un paziente in modo dinamico, con l'aiuto di un accessorio di ri-triage, le cui condizioni cliniche si sono deteriorate durante l'attesa di cercare un trattamento o quando il tempo di attesa assegnato è stato superato.

L'agente gestore code di trattamento gestisce il paziente nella coda di trattamento.

L'agente chiamante del trattamento chiama il paziente successivo in coda utilizzando l'ID dell'agente paziente fornito dal gestore della coda di trattamento.

Finalmente, l'agente medico richiede al chiamante del trattamento di chiamare un paziente per il trattamento. Il medico "trattiene" il paziente per una durata di tempo specificata allo scopo di "dare il trattamento". Una volta completato, il medico segnala il chiamante del trattamento per un altro paziente.

Quando arriva un paziente, viene creata una società di agenti composta da tutti i tipi sopraindicati.³⁷ Nella prima fase, il sistema prepara un elenco prioritario di pazienti presenti in pronto soccorso. Nella seconda fase, all'arrivo di un nuovo paziente, viene creato il suo agente e il sistema riordina l'elenco dei servizi prioritari. Nella terza fase, il sistema trova il personale più adatto e disponibile (Medico, Infermiere) per l'assegnazione del lavoro calcolando le proprie competenze e assegna il lavoro.

³⁶ Rahmat, Mohammad Hafidz, et al. "Agent-Based Modelling and Simulation of Emergency Department Re-Triage." Business Engineering and Industrial Applications Colloquium (BEIAC), 2013, <https://doi.org/10.1109/beiac.2013.6560119>.

³⁷ Iqbal, Sajid, et al. "Application of Intelligent Agents in Health-Care: Review." Artificial Intelligence Review, 2016 vol. 46, n. 1, pp. 83–112., <https://doi.org/10.1007/s10462-016-9457-y>.

I compiti svolti dal personale sono divisi in due categorie: compiti di Azione di Direzione (questi compiti richiedono all'attore medico di agire direttamente sul paziente, ad esempio farmaci, test) e Compiti di Azione Indiretta (non è richiesta alcuna azione del personale di direzione, ad esempio rapporti, monitoraggio).³⁸

Molto interessante appare anche la tecnologia agente per l'assistenza sanitaria dei pazienti affetti da Alzheimer.

In questo sistema di monitoraggio dei pazienti affetti da Alzheimer, gli agenti presentano un altro approccio per la programmazione delle risorse umane (medici, infermieri e assistenti) nei tempi di esecuzione e nell'allocazione ai pazienti nei reparti all'aperto di un ospedale.

Sul punto si riporta un progetto finanziato dall'UE, il DEM@CARE³⁹, che ha sviluppato una soluzione di gestione a ciclo chiuso per le persone che si trovano nelle fasi iniziali o poco gravi di demenza, attraverso il monitoraggio remoto multi-parametrico e l'analisi personalizzata a livello individuale delle misurazioni fisiologiche, comportamentali e dello stile di vita.

Il sistema di monitoraggio ambientale multisensore registra l'utilizzo dei servizi pubblici e dell'ambiente, profilando comportamenti e stili di vita, creando routine quotidiane.

Algoritmi di rilevamento visivo appositamente sviluppati, attraverso telecamere statiche e portatili, supportano il rilevamento di eventi in tempo reale.

L'analisi avanzata del rilevamento audio, invece, estrae caratteristiche quantitative significative che indicano lo stato comportamentale e mentale di ogni persona.

Utilizzando osservazioni oggettive, i dati del sistema forniscono ai medici un quadro dello stato di salute della persona con demenza e li avvisano della presenza di tendenze negative, in modo che possano essere prese le necessarie misure di risposta.

L'interfaccia paziente supporta funzioni semplici, come la gestione dei messaggi inviati al medico per organizzare gli appuntamenti. Ad esempio, l'utente riceve informazioni sulla durata del sonno piuttosto che sulla qualità e per evitare stress e scoraggiamento.

Il ciclo di comunicazione si conclude con le discussioni tra medici e i pazienti o con le persone che li assistono, in modo che possano fornire un supporto personalizzato.

³⁹ Dem@Care aspira a contribuire alla diagnosi tempestiva, alla valutazione, al mantenimento e alla promozione dell'auto-indipendenza delle persone con demenza, approfondendo la comprensione di come la malattia influisce sulla loro vita quotidiana e sul loro comportamento.

Implementa una soluzione di gestione remota multiparametrica a circuito chiuso che offre un feedback adattivo alla persona con demenza, includendo allo stesso tempo i medici nel follow-up remoto, consentendo loro di mantenere una visione completa dello stato di salute e dei progressi della persona interessata.

Un altro agente di notevole considerazione è il cosiddetto agente AGALZ⁴⁰ che viene utilizzato per pianificare la comunicazione e la conservazione dei registri dei pazienti.

Un altro interessante sistema che si occupa della coda del servizio di emergenza di un pronto soccorso è AMSEDR (Agent based Modelling and Simulation of Emergency Department Re-tirage), un lavoro di IA basato sulla simulazione che rivaluta i pazienti presenti.

Il sistema propone una valutazione in tempo reale per riorganizzare la coda in base alle condizioni dei pazienti e il deterioramento delle condizioni per gestire l'allocazione ottimale delle risorse (medici e infermieri) al compito più urgente.⁴¹

E' chiaro che i lavori di ricerca riportati nella letteratura e riguardanti l'applicazione delle tecniche della ricerca operativa nel settore della sanità, si sono concentrati generalmente su servizi ben particolari, servizi di emergenza come il pronto soccorso.

Tuttavia non manca strumentazione di intelligenza artificiale utilizzate nelle sala operatorie che sono uno dei servizi più costosi dell'ospedale.

La letteratura è piena anche di problemi che illustrano la complessità della gestione di una sala operatoria, sia in materia di dimensionamento, di pianificazione che di programmazione. Inoltre, la maggior parte dei lavori si basa su un approccio incentrato su una sola risorsa, spesso considerata critica, quali i letti di ricovero, i letti di rianimazione, la sala operatoria e gli infermieri.

In alcuni strumenti di IA⁴² si è proposto una pianificazione settimanale della sala operatoria dove le fasce orarie vengano riservate ai chirurghi e non per specialità dell'intervento.

L'algoritmo ha l'obiettivo di minimizzare l'uso delle sale operatorie, il costo degli straordinari orari di apertura del blocco e il tempo di attesa tra due interventi chirurgici.

Il problema della pianificazione viene risolto specificando la data dell'intervento per ogni paziente in base alla disponibilità dei chirurghi e delle sale operatorie. Il programma giornaliero è progettato per determinare la sequenza di operazioni in ogni sala operatoria per ogni giorno tenendo conto della disponibilità di letti di sveglia.

Gli esempi dell'utilizzo degli agenti intelligenti in medicina potrebbero essere tantissimi ed abbracciano tutto il settore sanitario.

⁴⁰ De Paz, Juan F., et al. "An integrated system for helping disabled and dependent people: AGALZ, AZTECA, and MOVI-MAS projects." Springer, Cham, 2015.

⁴¹ Noura Benhajji. "Système multi-agents de pilotage réactif des parcours patients au sein des systèmes hospitaliers. Gestion et management", Université de Lorraine, 2017

⁴² Fei, H., et al. "A Planning and Scheduling Problem for an Operating Theatre Using an Open Scheduling Strategy" Computers & Industrial Engineering, 2010 vol. 58, n. 2, pp. 221-230., <https://doi.org/10.1016/j.cie.2009.02.012>.

I casi riportati sottolineano come la ricerca scientifica tenda a favorire metodiche e procedimenti completamente automatizzati, al fine di evitare possibili bias da parte degli utenti.⁴³

⁴³ Pasceri, G. *“Intelligenza Artificiale, Algoritmo E Machine Learning: La Responsabilità Del Medico e Dell'amministrazione Sanitaria”*, Giuffrè Francis Lefebvre, 2021.

CAPITOLO II

APPLICAZIONE E PRINCIPALI FONTI NORMATIVE DELL’I.A. NELL’AMBITO SANITARIO

1. Qualificazione dei sistemi di I.A. in sanità.....pag. 34
2. Questioni etiche e giuridiche e rischi dell’I. A.....pag. 42
3. Il rischio del “deskilling” professionale in ambito sanitario.....pag. 50
4. Il ruolo dell’Europa in riferimento all’utilizzo dell’ I.A. – orientamenti per una I.A. affidabile.....pag. 55
5. Il GDPR e il diritto alla spiegazione.....pag. 62
6. Il Regolamento della Commissione europea sull’ I.A.: l’ Artificial Intelligence Act e il modello basato sul rischio.....pag. 75
7. OMS: Raccomandazioni e linee guida in merito allo sviluppo, all’applicazione e all’uso delle tecnologie di intelligenza artificiale (Artificial Intelligence, AI) nel settore della medicina.....pag. 82
8. Le sfide etiche e legali della intelligenza artificiale sanitaria negli USA.....pag. 88

1. Qualificazione dei sistemi di I.A. in sanità

In ambito sanitario esiste una varietà di tecnologie e di sistemi di I.A. che sono progettati e sviluppati per svolgere compiti diversi tra di loro, abbracciando i diversi campi della medicina. Basti pensare alle nuove tecniche sviluppate nelle operazioni chirurgiche che elevano le possibilità di buona riuscita degli interventi, oppure agli strumenti utilizzati per supportare le decisioni in campo diagnostico e terapeutico, o ancora per formulare programmi di prevenzione delle malattie o di riabilitazione.

Data l'eterogeneità della materia, è opportuno ricorrere ad una classificazione che permetta di evidenziare le caratteristiche dei diversi sistemi di intelligenza artificiale utilizzate in campo medico. La classificazione più agevole da fare è quella riconducibile alle diverse aree cliniche in cui si applicano i sistemi di robotica di I.A.: area assistenziale, area clinica, area riabilitativa.⁴⁴

La medicina preventiva e predittiva sono entrambe incluse nei programmi di assistenza sanitaria preventiva delle autorità pubbliche e sono rivolte a gruppi di persone che non sono malate (ancora). In campo assistenziale in Europa ci sono programmi di screening per monitorare la salute dei giovani, il cancro al seno, i problemi dentali ecc. e tipi di programmi di educazione pubblica diretti a gruppi specifici e aspetti specifici come l'esercizio fisico.

L'assistenza sanitaria preventiva è resa più specifica con l'assistenza sanitaria predittiva. L'assistenza sanitaria predittiva si basa sull'identificazione di gruppi ad alto rischio sulla base di statistiche che collegano i problemi di salute al background familiare, allo stile di vita, alla zona di vita, alle condizioni di lavoro, ai precedenti problemi di salute, ecc.

Il riconoscimento precoce dei gruppi a rischio ha bisogno di follow-up che comportano monitoraggio, screening ed eventualmente diagnosi. La robotica può avere un ruolo in queste attività evolvendosi gradualmente da applicazioni tradizionali per ampie fasce della popolazione ad applicazioni specificamente progettate per la valutazione individuale del paziente.

I sistemi diagnostici specifici saranno strettamente correlati ai dispositivi di trattamento che consentono il monitoraggio dei progressi del paziente.

Dal punto di vista della prevenzione, si possono individuare due possibili aree di innovazione. Il primo è la connessione delle funzionalità diagnostiche ai robot già utilizzati. Questi possono essere robot utilizzati per l'analisi del movimento o della coordinazione o robot di fitness intelligenti. L'altro gruppo di aree di innovazione sono i robot diagnostici dedicati, tra cui robot teleradiologici e robot miniaturizzati per la diagnosi interna.

⁴⁴ Pasceri, G. *“Intelligenza Artificiale, Algoritmo E Machine Learning: La Responsabilità Del Medico e Dell'amministrazione Sanitaria”* Giuffrè Francis Lefebvre, 2021.

Nell'analisi del movimento e della coordinazione, le sequenze di movimenti devono essere eseguite dai pazienti per scansionare le loro capacità motorie. Questi movimenti devono essere registrati e analizzati, facilitando la descrizione precisa e la valutazione delle capacità del paziente. Utilizzando tali procedure, è possibile identificare importanti disfunzioni muscolari e di coordinazione. L'uso di sofisticati programmi di screening per il sistema motorio e di coordinazione porta a informazioni dettagliate sull'insorgenza e lo sviluppo di malattie.

In questo contesto, i robot possono eseguire test complessi che sono noiosi e fastidiosi per i professionisti paramedici, in modi molto efficienti e riproducibili. Questa area di applicazione sta emergendo nel campo della robotica come spin-off dei sistemi robotici sviluppati per scopi terapeutici.

Un'altra area di applicazione strettamente correlata alla precedente area di innovazione potrebbe essere l'uso di sofisticate macchine per il fitness per l'intera popolazione cittadina.

La promozione dell'esercizio fisico è diventata un obiettivo molto importante nei programmi di salute pubblica a causa della relazione con una serie di "malattie da ricchezza" (obesità, diabete). Tuttavia, l'adattamento individuale alla situazione personale e al progresso della forma fisica deve ancora essere migliorato. Rafforzare la conformità o la motivazione a perseverare per periodi di tempo più lunghi offrendo programmi di allenamento più variabili e attraenti è anche uno dei fattori chiave per migliorare il successo dei programmi di fitness.

Per quanto riguarda l'assistenza professionale, è noto che questa sia un'importante istituzione sociale, con una forza lavoro consistente, ma si teme che il crescente invecchiamento della società porti a una pressione sociale troppo ampia.

La crescente domanda di servizi di assistenza e la crescente pressione sull'offerta di tali servizi, indica un futuro squilibrio tra domanda e offerta, non da ultimo a causa dell'aumento dell'età degli stessi prestatori di assistenza professionale. Si prevedono gravi carenze nelle professioni infermieristiche e assistenziali, mentre l'immagine che il settore ha nel mercato del lavoro si è già deteriorata negli ultimi decenni. Sebbene vi sia una crescente tendenza a fornire assistenza ai pazienti che vivono a casa il più a lungo possibile, molta assistenza è e sarà fornita in istituti di assistenza come ospedali, case di riposo e case di cura. Le attività dei fornitori di assistenza professionale che possono essere supportate dai robot includono il trasferimento e il sollevamento dei pazienti, i processi logistici, la fornitura di assistenza nelle attività ADL: igiene, nutrizione del paziente, monitoraggio dei pazienti, ecc.

L'introduzione della tecnologia nell'assistenza, compresa la robotica, è considerata come una potenziale soluzione per il problema dell'invecchiamento.

Ancora si pensi all'aiuto logistico robotizzato per infermieri. Gran parte delle attività di assistenza professionale è di natura logistica. Comprende sia la soddisfazione delle esigenze quotidiane dei pazienti (distribuzione di posta o prodotti per la cura personale, attività di pulizia ambientale), sia la fornitura di farmaci e cibo ai pazienti.

Tradizionalmente, questo lavoro viene svolto manualmente, senza molta automazione. Lo sviluppo di sistemi intelligenti consente l'uso di robot per automatizzare la consegna di medicinali e alimenti. Ciò può migliorare la sicurezza, la conformità e l'affidabilità, riducendo al contempo i costi di manodopera.

E' vero che molte attività logistiche non legate ai pazienti negli ospedali possono già essere eseguite da sistemi robotizzati. Sebbene sia già disponibile la tecnologia necessaria, si richiederà l'adattamento di quest'ultima all'ambiente del paziente per essere efficace in questo settore di innovazione di IA. Anche qui, la maggior parte dei principali driver di innovazione della robotica giocano un ruolo importante (miniaturizzazione, intelligenza attraverso le TIC, materiali ad alte prestazioni).

In generale, gli sviluppi in questo campo si basano su tecnologie più generali utilizzate anche in altri settori. I sistemi dedicati ad alta tecnologia per gli ospedali sono ancora in fase di sviluppo.

Fare una diagnosi di solito richiede la presenza di un medico.⁴⁵

I nuovi sviluppi della robotica in materia assistenziale rendono possibile il monitoraggio da remoto dei pazienti. Questi sistemi possono includere funzionalità di comunicazione, sistemi sensoriali, sistemi di gestione della conoscenza e strumenti intelligenti di interfaccia uomo-macchina (HMI).

Gli sviluppi nei sistemi di telediagnostica o monitoraggio e di essere suddivisi nei seguenti concetti di innovazione:

- Diagnostica avanzata, sistemi di sensori automatizzati o sistemi di imaging che possono essere tele-operati o presentano un livello di automazione;
- Il paziente può utilizzare la consultazione remota attraverso un robot che include funzionalità di comunicazione e sistemi sensoriali limitati per l'esame del paziente;
- Monitoraggio del paziente, sistemi robotici che monitorano i segni vitali e altre caratteristiche dei pazienti e segnalano incidenti o trasmettono informazioni continue.

In generale, quest'area di applicazione è in fase di sviluppo, ma con un numero limitato di interessanti applicazioni commerciali disponibili. Tuttavia, si assiste a molti sviluppi per aumentare ulteriormente l'intelligenza dei suddetti sistemi.

⁴⁵ Schönberger D., "Artificial intelligence in healthcare: a critical analysis of the legal and ethical implications", *International Journal of Law and Information Technology*, 2017 V. 27, pp. 171–203, <https://doi.org/10.1093/ijlit/eaz004>

Sono ancora numerosi gli esempi dell'utilizzo della robotica in medicina.

Nell'ambito dell'area clinica le tecnologie più rilevanti riguardano i sistemi a supporto alle decisioni diagnostiche e terapeutiche e alla chirurgia robotica.

Si pensi che in genere la diagnosi delle malattie viene spesso eseguita da un endoscopio oppure utilizzando strumenti sofisticati che possono includere la funzionalità di prelevare campioni di tessuto.

Queste procedure sono considerate onerose per il paziente. L'uso della robotica, combinato con l'ulteriore miniaturizzazione dei sensori, ha portato allo sviluppo di microcapsule endoscopiche in miniatura e wireless. Queste capsule mediche intelligenti sono state sviluppate per la prima volta nel 2001 come prodotto per l'ispezione delle malattie del tratto digestivo. È stato il primo promettente dispositivo in miniatura non invasivo nello sviluppo verso una vera microcapsula robotica.

Consentire al medico di fermare, spostare e orientare la capsula a comando migliorerebbe le capacità diagnostiche consentendo all'operatore di dare un'occhiata più da vicino o tornare indietro per riesaminare una regione di interesse. Inoltre, una volta ottenuto un controllo affidabile e aumentate le capacità del dispositivo, le micro-capsule endoscopiche aggiungono il controllo del movimento e le telecamere agli attuali endoscopi della capsula.

Il controllo teleoperato e semiautonomo di queste capsule è una nuova sfida robotica ed è un argomento che sarà ripreso successivamente in questo elaborato di tesi per individuare l'attribuzione delle responsabilità in caso di danno cagionato dalla diagnosi o dall'intervento sbagliato.

L'utilizzo dell'I.A. negli ospedali sicuramente più diffuso e di facile individuazione è quello legato alla chirurgia robotica.

Una delle principali attività del sistema sanitario è l'esecuzione di interventi medici. Ciò include la chirurgia, ma anche interventi minori come fare una biopsia. La ricerca nel campo della robotica a supporto dell'intervento medico è molto attiva oggi. In quest'area, i robot utilizzati per supportare il chirurgo stanno iniziando a prendersi in carico alcuni compiti del medico.

Il supporto robotico dell'intervento medico può aiutare a svolgere compiti che i medici da soli non sarebbero in grado di svolgere a causa della precisione, della resistenza e della ripetibilità dei sistemi robotici. Inoltre, i robot sono ben in grado di operare in uno spazio contenuto all'interno del corpo umano. Per questi motivi i sistemi robotici sono considerati un fattore importante nel futuro della chirurgia. In particolare, i robot stanno contribuendo al miglioramento della chirurgia non invasiva o minimamente invasiva e a rendere migliori i risultati della chirurgia. Con i sistemi chirurgici robotici è possibile eseguire chirurgia a cielo aperto, chirurgia mini-invasiva, telechirurgia a distanza, pianificazione preoperatoria, prove chirurgiche, navigazione intraoperatoria (chirurgia guidata da immagini) e simulazione chirurgica da un unico posto.

Oggi, i robot sono effettivamente in uso nella chirurgia del cervello, del cuore, del midollo spinale, del ginocchio, della gola, dei reni e degli occhi. Poiché il sistema chirurgico robotico migliora la coerenza e la qualità, sta diventando sempre più popolare. In questo ambito si possono distinguere le seguenti aree di innovazione:

- Microchirurgia assistita da robot: qui i chirurghi possono essere assistiti da robot per la precisione e la microchirurgia. Le macchine possono ridimensionare i movimenti del chirurgo, filtrare il tremore della mano e impedire al chirurgo di tagliare accidentalmente lì dove non dovrebbe. La pratica medica pertinente comprende interventi chirurgici agli occhi, all'orecchio, al naso, alla gola, al viso, alle mani e ai crani. In sala operatoria i robot possono aiutare il chirurgo a passare gli strumenti e le forniture sterili. Il sistema può anche fornire al chirurgo uno o più bracci in più per tenere qualcosa: l'arto di un paziente, uno strumento o una macchina fotografica. Questo braccio può essere controllato attraverso pedali e rimane perfettamente fermo.
- Chirurgia di precisione robotizzata: alcune procedure in chirurgia sono adatte per l'assistenza robotica. Queste procedure sono ben definite e possono essere pianificate in anticipo in un modello 3D del paziente. Così un robot può lavorare automaticamente e più facilmente grazie alla sua natura statica e predittiva. Questi robot sono molto utili per l'esecuzione di compiti pre-pianificati in materiali duri come le ossa.
- Dispositivi robotici per la chirurgia mini-invasiva: la chirurgia mini-invasiva (MIS) sta diventando sempre più popolare perché causa meno traumi per il paziente che si riprendono più velocemente. Per il chirurgo il MIS è spesso più impegnativo e complicato perché tutto deve essere fatto attraverso alcuni piccoli fori.

La chirurgia MIS spesso richiede più tempo, il chirurgo non può vedere direttamente su cosa sta lavorando (deve guardare attraverso un endoscopio o su un display), non può sentire cosa sta facendo e l'uso degli strumenti è meno intuitivo.

La chirurgia mini-invasiva consente solo agli strumenti e non alle mani del chirurgo di entrare nel corpo del paziente. Rendere possibile questo tipo di chirurgia richiede lo sviluppo di strumenti e strumenti specializzati. Telecamere, meccanica, supporto all'orientamento e sistemi di imaging intraoperatorio sono esempi. Alcuni di questi sistemi soddisfano la definizione di robot.

- Micro e nanobot medici: si fa riferimento a nanorobot con scopi diagnostici e di screening, sistemi di somministrazione di farmaci e applicazioni di terapia genica. I sistemi di nanobot sono ancora in fase di ricerca e di sviluppo ma il loro potenziale è evidente. Alcuni sistemi molto interessanti permettono l'identificazione e la distruzione di cellule tumorali.
- Chirurgia remota: Alcuni interventi chirurgici sono complicati e richiedono chirurghi altamente qualificati. Questi esperti a volte non sono disponibili in un ospedale, o anche in un determinato

territorio. La chirurgia remota può consentire la possibilità di queste procedure utilizzando un sistema robotico. L'operatore umano con l'ausilio di un monitor esegue le manovre necessarie che vengono teletrasmesse e ripetute su un paziente da un sistema robotico con estrema precisione. Il chirurgo e i robot sono dotati di sensori come telecamere, microfoni e sensori pressori che permettono di riprodurre fedelmente gli stimoli sensoriali.

- Assistenza robotizzata per piccoli interventi medici: oltre alla chirurgia, si svolgono anche altri piccoli interventi medici. Questi possono essere fatti da un infermiere qualificato o da medici. Esempi sono il prelievo di campioni di sangue, le biopsie e l'esecuzione di punture. Trattandosi normalmente di piccoli interventi più o meno routinarie, spesso possono essere robotizzati. Sono in fase di sviluppo effetti finali robotici per facilitare le procedure mininvasive basate su ago guidate da immagini, come l'ablazione del tumore, la biopsia, la toracentesi e il prelievo di sangue. Un sistema robotico può prelevare campioni di sangue autonomamente dall'avambraccio.
- Assistenza chirurgica robotizzata per uso diagnostico: negli ultimi anni, i sistemi di informazione e raccolta si sono sviluppati nel campo della medicina. Quest'ultimo è considerato come uno dei campi di applicazione dell'intelligenza artificiale, per le caratteristiche specifiche che il sapere medico ha in relazione alle decisioni cliniche. La ricerca ha portato allo sviluppo di sistemi di supporto alle decisioni di vario tipo e complessità. I più sofisticati sono senza dubbio i sistemi attraverso i quali il sapere medico, opportunamente organizzato, deve raggiungere obiettivi diagnostici nelle varie discipline, secondo strategie definite da esperti. Una proprietà caratteristica dei sistemi esperti è la capacità di fornire spiegazioni esaustive dei ragionamenti, delle procedure e delle conclusioni a cui sono pervenuti. Sono uno strumento ottimale per assistere la decisione clinica.⁴⁶

La terza area applicativa dell'utilizzo della robotica in medicina riguarda l'area riabilitativa che si concreta prettamente sul trattamento dei pazienti con disabilità fisiche e/o mentali.

La robotica riabilitativa può essere utilizzata sia in ambiente sanitario che domestico. Le malattie principali per le quali viene impiegata riguardano principalmente la riabilitazione motoria e la riabilitazione delle malattie cognitive.

Nel primo caso sono state sviluppate tecnologie robotiche per la riabilitazione di arti superiori e inferiori. Tali sistemi permettono un movimento guidato degli arti e permettono al paziente di acquisire forza. Nel secondo caso i sistemi di I.A. sono progettati per interagire con gli esseri umani,

⁴⁶ Lagioia, F. *“L'intelligenza Artificiale in Sanità: Un'analisi Giuridica”*, Giappichelli, 2020.

infatti nella maggior parte dei casi questi sistemi vengono considerati come oggetti o animali domestici il cui potenziale terapeutico è ampiamente riconosciuto.⁴⁷

Ciò che è evidente dall'esame dei sistemi di I.A. utilizzate o di prossima introduzione nel settore sanitario è la difficoltà di raggruppare tutte in una stessa categoria concettuale, capace di caratterizzarne in modo univoco. Queste tecnologie hanno proprietà estremamente eterogenee e diverse, sebbene siano accomunate da alti livelli di automazione.

Alcuni possono certamente essere inclusi nella nozione di sistema dotato di alcune funzioni essenziali, come la capacità di agire su stimoli ambientali con rilevamento e ragionamento logico, il principale dei quali è l'automazione, a volte anche solo parziale della fisica, come un sistema robotico per la semplice distribuzione di farmaci.⁴⁸

La classificazione di questi sistemi diventa rilevante per gli operatori del diritto. Infatti, la classificazione, oltre a individuare le evidenze tecniche deve fornire necessariamente indicazioni giuridiche, per individuare, in maniera correlata, i criteri di responsabilità in caso di danni agli individui.

E' possibile individuare diversi criteri di classificazione. Il primo criterio fa riferimento alle aree applicative: è intuibile che questo criterio faccia riferimento alle diverse aree in cui opera la medicina. Il criterio permette di verificare le responsabilità e i profili giuridici contestualizzando il campo in cui operano le tecnologie coinvolte.

Il secondo criterio considera l'ambiente e il contesto in cui un sistema di IA è utilizzato. È possibile distinguere, ad esempio, un ambiente ospedaliero o domestico e un ambiente strettamente o non controllato.

Nel caso in cui il sistema di IA operi in un ambiente controllato, l'accesso sarà concesso solo a personale autorizzato che potrebbe essere stato addestrato all'uso di tali sistemi, mentre nel caso di un ambiente non controllato ambiente, il sistema si sposterà in aree in cui l'accesso è riservato a coloro che non hanno competenze specialistiche come pazienti, familiari in visita, ecc.

Rispetto al contesto è possibile effettuare un'ulteriore distinzione tra tecnologie per l'assistenza professionale e quelle per l'assistenza personale.

Nel primo caso il sistema è utilizzato da operatori specializzati, autorizzati, addestrati, preposti all'avviamento, al monitoraggio, all'interruzione le funzioni del sistema di intelligenza artificiale, come ad esempio il personale. Questa ipotesi include sistemi di supporto alle decisioni cliniche, robot

⁴⁷ Lagioia, F. "L'intelligenza Artificiale in Sanità: Un'analisi Giuridica" Giappichelli, 2020.

⁴⁸ Lagioia, F., et al. "Automated detection of unfair clauses in online consumer contracts." Legal Knowledge and Information Systems: JURIX 2017: The Thirtieth Annual Conference, 2017, Vol. 302 p. 145.

utilizzati per la consegna di cibo e altri negli ospedali, sistemi robotici per sistemi di chirurgia robotica riabilitativa.

Diversamente, i sistemi di assistenza personale possono essere utilizzati da persone prive di conoscenze professionali e specialistiche. Rientrano in questa categoria, ad esempio, alcuni robot di assistenza domiciliare e i sistemi di assistenza domestica.⁴⁹ Si pensi, ad esempio, ai sistemi utilizzati per supporto alla deambulazione.

Un ulteriore criterio si riferisce alla natura fisica o intangibile di un sistema. La distinzione può essere fatta tra robot e software.

La natura del sistema sarà necessariamente correlata al tipo di interazione tra gli operatori sanitari e lo strumento tecnologico.

Per l'individuazione dei criteri di classificazione dei sistemi di intelligenza artificiale rileva senza dubbio l'interazione tra uomo e macchina. Questo tipo di criterio diventa particolarmente rilevante perché configurerà anche il tipo di rischio e il potenziale danno per coloro che vengono sottoposti a queste tecnologie.

Rilevante è anche il criterio che riguarda i diversi livelli di automazione del sistema rispetto alle diverse funzioni cognitive a esso affidate.

Chiaramente l'affidamento che l'operatore può avere nei confronti della macchina rileverà ai fini dell'individuazione di responsabilità per eventuali danni che possono essere cagionati dall'utilizzo della stessa.

La attuale letteratura e i riferimenti bibliografici individuano diverse definizioni e poco chiare del concetto di automazione, rispecchiando l'eterogeneità del processo tecnologico.

Ultimo criterio è quello della trasparenza e della capacità dei sistemi di I.A. di fornire spiegazioni chiare rispetto al proprio funzionamento e ai risultati dei calcoli algoritmici.

Questi criteri, come già accennato, diventeranno rilevanti ai fini dell'individuazione delle responsabilità per i danni cagionati alla salute dei pazienti e verranno approfonditi nei prossimi paragrafi di questo elaborato di tesi.

2. Questioni etiche e giuridiche e rischi dell'I. A

Il connubio sempre più stretto fra tecnologia e medicina rende necessaria una disamina sul piano sia etico che giuridico.

Come già ampiamente evidenziato i Big data sono un fenomeno complesso con capacità di archiviazione, di velocità computazionale, che permettono a basso costo la raccolta dei dati e l'analisi predittiva.

Questi dati, che vengono continuamente raccolti, sono infinitamente collegati in rete e altamente flessibili e possono essere analizzati in contesti molto disparati.

L'intelligenza artificiale (IA) può essere considerata come una chiave per sbloccare il valore dei big data mentre l'apprendimento automatico è uno dei meccanismi tecnici che sostiene e facilita l'IA.

È evidente che l'uso della profilazione e del processo decisionale automatizzato reso possibile dall'apprendimento automatico e dall'intelligenza artificiale dovrà essere sottoposto ad una disciplina di tutela più rigorosa rispetto alle protezioni della privacy già note e richiederà una supervisione etica.

I big data hanno sollevato questioni in riferimento ai loro metodi di ricerca sulla scienza dei dati e i paradigmi etici di ricerca ai quali bisogna attenersi.

L'etica della ricerca tradizionale basata su presupposti concettuali, normativi e istituzionali dei dati relativi all'uomo non è facilmente applicabile all'analisi dei big data. I conflitti che ne derivano hanno accelerato un esame tempestivo della privacy e delle considerazioni etiche nella raccolta e nell'analisi dei big data.

L'adozione di un approccio etico all'elaborazione dei big data è guidata da due fattori principali.

Nel settore pubblico, l'evidenza di una mancanza di consapevolezza pubblica sull'uso dei dati e sulla portata della condivisione dei dati ha portato a richieste di rendere esplicite le politiche etiche.

L'imperativo commerciale nel settore privato, invece, è quello di mitigare il rischio di danni alla reputazione dovuti alla sfiducia pubblica e alla svalutazione del marchio.⁵⁰

E' chiaro che la legislazione sulla privacy non è sufficiente, per questo motivo i quadri etici per l'analisi e la ricerca dei big data sono altamente contestati e in continua evoluzione.

Al centro del dibattito etico ci sono le conseguenze della velocità, della capacità e della continua generazione di big data, nonché un cambiamento nella relazionalità, flessibilità, riutilizzo e decontestualizzazione dei dati.

⁵⁰ Belk, Russell. "Ethical issues in service robotics and artificial intelligence." *The Service Industries Journal*, 2021 pp. 860-876.

In particolare, i problemi si intensificano se si pensi che con l'intensificazione della profilazione algoritmica gli individui non sono trattati come persone ma come aggregati temporanei di dati trattati su scala industriale. Ma gli esseri umani non sono tutti identici e non possono essere considerati come dati. È necessario, infatti, considerare i valori umani che devono essere compresi e implementati all'interno di un contesto sociale, culturale, politico, economico e tecnologico in cui vengono realizzati i dati personali e l'esperienza personale.⁵¹

Pertanto, l'etica digitale dovrebbe tenere conto del rapporto in ampia evoluzione tra realtà digitali e umane. I big data generano nuove domande etiche su cosa significhi essere umani in relazione ai dati, sulla conoscenza umana e sulla natura dell'esperienza umana.

Dal punto di vista dell'Unione Europea, High-Level Expert Group on AI, istituito dalla Commissione europea, ha elaborato un documento che offre un quadro etico digitale che fornirà nuovi termini per identificare, analizzare e comunicare nuove realtà umane, al fine di sostituire le tradizionali questioni basate sui valori e identificare nuove sfide in vista dei valori in gioco e dei cambiamenti tecnologici esistenti e prevedibili.

Lo scopo dell'etica digitale non è solo quello di rendere conto del presente, ma anche di svolgere una funzione di previsione. Il cambiamento è duplice. In primo luogo, l'oggetto della regolamentazione legale (cioè un individuo) può diventare meno interessante, come fenomeno nel qui e ora e più oggetto di speculazioni ragionate sul suo ruolo futuro, il tutto basato sui poteri predittivi dei big data e dell'elaborazione algoritmica. In secondo luogo, mentre l'analisi delle questioni giuridiche viene spinta verso il futuro, ciò che è inteso come esistente nel futuro viene attirato nelle valutazioni del presente. Ad esempio, le stime di ciò che il futuro riserverà, generate attraverso i modelli raccolti nell'analisi dei big data, stanno continuamente guadagnando importanza per il modo in cui la giustizia penale opera oggi e pretende di operare in futuro.

L'attenzione dell'etica digitale è principalmente su questioni meta-etiche e considera domande più generali e fondamentali su cosa significhi fare affermazioni sull'etica e sulla condotta umana nell'era digitale, quando le condizioni di base della "umanità" sono sotto la pressione dell'interconnettività, del processo decisionale algoritmico, dell'apprendimento automatico, della sorveglianza digitale e dell'enorme raccolta di dati personali, su ciò che può e deve essere mantenuto e ciò che può e deve essere adattato, dall'etica normativa tradizionale. I seguenti esempi forniscono informazioni sulla necessità di un quadro etico digitale:

⁵¹ Scalzini, Silvia. "Alcune questioni a proposito di algoritmi, dati, etica e ricerca." *Rivista italiana di medicina legale e del diritto in campo sanitario* 2019 pp. 169-178.

1. Dall'individuo al soggetto digitale: i dati non esauriscono né l'identità personale, né le qualità delle comunità a cui appartengono gli individui, e la protezione dei dati non riguarda solo la protezione dei dati, ma principalmente la protezione delle persone dietro i dati. La questione è se la rappresentazione digitale delle persone possa esporle a nuove forme di vulnerabilità e danno.
2. Dalla vita analogica a quella digitale: il governante e il governato sono distinti ma legati da principi reciprocamente riconosciuti di obbligo legale e responsabilità. Le tecnologie digitali hanno cambiato questa situazione. L'uso di algoritmi e grandi set di dati può plasmare e dirigere la vita degli individui, quindi sempre più governata sulla base dei dati generati dai propri comportamenti e interazioni. La distinzione tra le forze che governano la vita quotidiana e le persone che sono governate al suo interno diventa così più difficile da discernere. Il comportamento può essere governato da "nudging", cioè da suggerimenti minuti e appena percettibili, che possono assumere una varietà di forme e che possono modificare l'ambito delle scelte che gli individui hanno o credono di avere.
3. Da una società del rischio a una società valutata: la valutazione del rischio viene effettuata utilizzando tecniche di calcolo della probabilità, consentendo di mettere in comune individui e identificare situazioni con lo stesso livello di rischi tra loro per la Big Data Privacy e le sfide etiche. Nell'era digitale, gli algoritmi supportati dai big data possono fornire una comprensione molto più dettagliata e granulare dei comportamenti e delle propensioni individuali, consentendo valutazioni del rischio più individualizzate e la ripartizione dei costi effettivi per ogni individuo; tale valutazione del rischio minaccia i principi contrattuali o generali e le idee di solidarietà ampiamente condivise. In questa società con punteggio, gli individui possono essere iper-indicizzati e iper-quantificati. Le convinzioni e i giudizi su di loro possono essere fatti attraverso algoritmi di credito opachi o di punteggio sociale che devono essere aperti alla negoziazione o alla contestazione.
4. Dall'autonomia umana alla convergenza di esseri umani e macchine: un numero crescente di artefatti tecnologici, da protesi come occhiali e apparecchi acustici, a smartphone, GPS, occhiali a realtà aumentata e altro ancora, può essere vissuto in un rapporto simbiotico con il corpo umano. Questi artefatti sono vissuti meno come oggetti dell'ambiente che come un mezzo attraverso il quale l'ambiente viene vissuto e agito. Come tali, possono tendere verso un'inquadratura senza soluzione di continuità della nostra percezione della realtà. Possono modellare la nostra esperienza del mondo in modi che possono essere difficili da valutare criticamente.

Questo fenomeno di incorporazione o addirittura incarnazione di tecnologie è ancora più intenso ogni volta che i dispositivi vengono impiantati nel corpo. Una frontiera parallela di convergenza tra uomo e macchine è sul punto di essere attraversata da macchine intelligenti, o meglio "autonome", che sono in grado di adattare i loro comportamenti e piuttosto che semplicemente eseguire comandi umani, collaborare o addirittura sostituire agenti umani per aiutarli a identificare i problemi da risolvere, o per identificare i percorsi ottimali per trovare soluzioni a determinati problemi.

5. Dalla responsabilità individuale alla responsabilità distribuita: i problemi di molte mani e i problemi dell'azione collettiva e dell'inazione collettiva possono portare a tragedie dei beni comuni e valutazioni morali problematiche di complessi sforzi umani, sia a bassa che ad alta tecnologia, in cui un certo numero di persone agisce congiuntamente attraverso catene causali distanti, pur essendo separate nel tempo e nello spazio l'una dall'altra e dai risultati aggregati del loro libero arbitrio individuale. I problemi di allocazione e attribuzione delle responsabilità sono esacerbati dalla configurazione in rete del mondo digitalizzato.

6. Dalla giustizia penale alla giustizia preventiva: nella pratica legale, l'individuazione e l'indagine del crimine non è più solo una scienza degli atti criminali, dell'identificazione e del giudizio di eventi scritti da singoli attori identificabili e responsabili in condizioni precise e in termini di responsabilità morale e legale, ma anche un calcolo statisticamente supportato della probabilità di crimini futuri, una strutturazione del governo del crimine attorno alla scienza della possibile trasgressione e della possibile colpa, rimuovendo il carattere morale dall'equazione. L'obiettivo della giustizia penale rimane lo stesso: garantire la sicurezza all'interno della società aderendo allo stesso tempo a standard elevati in materia di diritti umani e Stato di diritto. Tuttavia, il cambiamento che segna uno dei principali sfondi dell'era digitale e richiede una nuova etica digitale è quello di cercare di prevedere in anticipo il comportamento criminale, utilizzando l'output di analisi basate sui big data e algoritmi intelligenti per guardare al futuro.⁵²

C'è urgenza nello sviluppo di un quadro etico digitale.

Il gruppo europeo sull'etica della scienza e delle nuove tecnologie ha tentato, inoltre, di definire un quadro etico e legale riconosciuto per la progettazione, produzione, uso e distribuzione di sistemi di intelligenza artificiale, robotica e autonomi, individuando una serie di "principi etici fondamentali", fondati sui valori su cui si fonda la Carta dei diritti fondamentali dell'Unione europea.

⁵² Maggino, Filomena, and Gabriele Cicerchia. "Algoritmi, etica e diritto" in *Diritto Dell'Informazione e Dell'Informatica (II)*, 2019, vol. 3, p. 1161.

Invece, al gruppo di esperti di intelligence ad alto livello nominato dalla Commissione europea è stato incaricato di fornire linee guida etiche per lo sviluppo di questi. Ed anche fuori dai confini europei si discute dei problemi etici, come dimostrano, ad esempio, le linee Guida della Japanese Society for Artificial Intelligence⁵³.

Nella sua dichiarazione su “Artificial Intelligence, Robotics and Autonomous”, il Gruppo europeo sull'etica nelle scienze e nelle nuove tecnologie ha infatti sottolineato la necessità che “i sistemi autonomi siano guidati da una genuina preoccupazione per la ricerca e il principio della 'responsabilità' della società da parte degli sviluppatori e della cooperazione accademica per la protezione dei diritti e dei valori fondamentali e l'obiettivo è il design volto a l'incorporazione per non diminuirli.⁵⁴

Nel documento "Ethics Guidelines for Trustworthy AI", che sarà meglio evidenziato nei prossimi paragrafi, gli esperti di intelligenza artificiale nominato dalla Commissione Europea descrivono il quadro di riferimento per supportare lo sviluppo dei sistemi di IA.

In particolare, il documento individua i diritti fondamentali⁵⁵, l'etica, i valori e alcuni principi operativi che guidano lo sviluppo, la regolazione e l'uso dell'intelligence presentandoli come cerchi concentrici, che costituiscono il "scopo etico", la guida per creare strumenti affidabili di intelligenza artificiale.

Il documento individua l'etica dell'IA come un sottocampo dell'etica applicata, che si concentra sulle questioni etiche sollevate dallo sviluppo, dalla distribuzione e dall'uso dell'IA.

La sua preoccupazione centrale è identificare come l'IA possa far progredire o sollevare preoccupazioni per la buona vita degli individui, sia in termini di qualità della vita, sia di autonomia umana e libertà necessarie per una società democratica.

Un'IA affidabile può migliorare la prosperità individuale e il benessere collettivo generando prosperità, creazione di valore e massimizzazione della ricchezza.

⁵³Ryan, Mark, and Bernd Carsten Stahl. "Artificial intelligence ethics guidelines for developers and users: clarifying their content and normative implications" in Journal of Information, Communication and Ethics in Society, 2020.

⁵⁴ Scalzini, S. "Alcune questioni a proposito di algoritmi, dati, etica e ricerca" in Rivista italiana di medicina legale e del diritto in campo sanitario, 2019, pp.169-178.

⁵⁵ Nel documento sono raggruppati nelle seguenti categorie: “Rispetto della dignità umana, Libertà della persona, Rispetto della democrazia, della giustizia e dello Stato di diritto, Uguaglianza, non discriminazione e solidarietà, compresi i diritti delle persone a rischio di esclusione, Diritti dei cittadini”.

Può contribuire al raggiungimento di una società equa, contribuendo ad aumentare la salute e il benessere dei cittadini in modo da promuovere la parità nella distribuzione delle opportunità economiche, sociali e politiche.

Come per qualsiasi tecnologia potente, l'uso dei sistemi di IA nella società solleva diverse sfide etiche, ad esempio relative al loro impatto sulle persone e sulla società, alle capacità decisionali e alla sicurezza. Se utilizzeremo sempre più l'assistenza o delegheremo le decisioni ai sistemi di IA, dobbiamo assicurarci che questi sistemi abbiano un impatto equo sulla vita delle persone, che siano in linea con valori che non dovrebbero essere compromessi e in grado di agire di conseguenza, e che adeguati processi di responsabilità possano garantirlo.

L'Europa deve definire quale visione normativa di un futuro immerso nell'IA vuole realizzare e capire quale nozione di IA dovrebbe essere studiata, sviluppata, implementata e utilizzata in Europa per raggiungere questa visione.

Il documento specifica che un codice etico specifico per dominio – per quanto coerenti, sviluppate e granulari possano essere le sue versioni future – non può mai funzionare come sostituto del ragionamento etico stesso, che deve sempre rimanere sensibile ai dettagli contestuali che non possono essere catturati nelle linee guida generali. Oltre a sviluppare una serie di regole, garantire un'IA affidabile ci impone di costruire e mantenere una cultura e una mentalità etiche attraverso il dibattito pubblico, l'educazione e l'apprendimento pratico.

Nell'UE, il Gruppo europeo per l'etica delle scienze e delle nuove tecnologie ("GEE") ha proposto una serie di 9 principi di base, basati sui valori fondamentali stabiliti nei trattati e nella Carta dell'UE.⁵⁶ Questi principi etici possono ispirare strumenti normativi nuovi e specifici, possono aiutare a interpretare i diritti fondamentali e possono guidare la logica per lo sviluppo, la diffusione e l'utilizzo dei sistemi di IA, adattandosi dinamicamente all'evoluzione della società stessa.

I sistemi di IA dovrebbero migliorare il benessere individuale e collettivo. Questi principi etici sono radicati nei diritti fondamentali e devono essere rispettati al fine di garantire che i sistemi di IA siano sviluppati, implementati e utilizzati in modo affidabile. Sono specificati come imperativi etici, in modo tale che i professionisti dell'IA dovrebbero sempre sforzarsi di aderirvi. Senza imporre una gerarchia, il documento elenca i principi correlati ai diritti fondamentali su cui si basa nella Carta dell'UE.

Questi sono i principi di:

⁵⁶ Quintarelli, S., et al. *“An Ethical Perspective on Artificial Intelligence: Principles, Rights and Recommendations/Una Prospettiva Etica Sull'intelligenza Artificiale: Principi, Diritti E Raccomandazioni”* in BIOLAW JOURNAL, 2022, consultabile al <https://hdl.handle.net/11311/1172731>.

- Rispetto dell'autonomia umana
- Prevenzione dei danni
- Equità
- Spiegabilità

Questi principi devono delineare le zone grigie tra il lecito e l'illecito e devono dunque essere precetti etici per funzionalizzare i comportamenti verso un fine sempre e comunque benefico per loro e per la società.

Sulla base di questi principi si dovrà ispirare anche lo sviluppo di protezioni legislative della privacy, che è la parte più spinosa delle problematiche etiche legate all' I.A.

Per consentire pratiche di tipo big data in generale, è necessario implementare quadri di privacy diffusi nuovi o modificati per enti pubblici e privati per proteggere la privacy degli individui e garantire un uso equo ed etico delle loro informazioni personali.

L'analisi dei big data è distintiva nella raccolta di quantità significative di dati, riproponendo l'uso di tali dati, utilizzando l'anonimizzazione nell'analisi, generando nuovi dati da queste analisi e avendo opacità nell'elaborazione dei dati.

L'Information Accountability Foundation⁵⁷ ha distinto quattro tipi di nuovi dati prodotti dall'analisi dei big data:

1. Fornire dati forniti consapevolmente da individui, ad esempio quando si compila un modulo online.
2. Dati osservati che vengono registrati automaticamente, ad esempio da cookie o sensori online o televisione a circuito chiuso (CCTV) collegata al riconoscimento facciale.
3. Dati derivati prodotti da altri dati in modo relativamente semplice e diretto, ad esempio calcolando la redditività del cliente dal numero di visite a un negozio e dagli articoli acquistati.
4. Dati dedotti basati su probabilità e prodotti utilizzando un metodo di analisi più complesso per trovare correlazioni tra i set di dati e utilizzarli per classificare o profilare individui e popolazioni, ad esempio calcolando punteggi di credito o prevedendo futuri risultati di salute.⁵⁸

⁵⁷ The Information Accountability Foundation (IAF) si concentra su complesse questioni di politica globale, promuovendo sistemi di governance delle informazioni efficaci e responsabili, facilitando nel contempo un ecosistema digitale affidabile che serva le persone e sia rispettoso del loro diritto fondamentale a un trattamento equo. Per approfondimento <https://informationaccountability.org>

⁵⁸ Mittelstadt, Brent Daniel, and Luciano Floridi. “*The Ethics of Big Data: Current and Foreseeable Issues in Biomedical Contexts*” in *Law, Governance and Technology Series*, 2016, pp. 445–480, https://doi.org/10.1007/978-3-319-33525-4_19

Pertanto, il principio della privacy della raccolta diretta da un individuo per uno scopo specifico è messo in discussione dai big data, che incidono sull'autonomia personale di un individuo basata sul suo diritto di controllare i propri dati personali e il trattamento di tali dati.

Il controllo richiede consapevolezza dell'uso dei dati personali e la reale libertà di scelta. Tali condizioni, essenziali per la tutela dei diritti fondamentali, ed in particolare del diritto alla protezione dei dati personali, possono essere soddisfatte attraverso diverse soluzioni giuridiche adattate in base al contesto sociale e tecnologico.

Sorge infatti la problematica relativa al consenso informato, significativo perché l'analisi dei big data coinvolge dati che possono essere una raccolta continua nel tempo e le conseguenze previste non sono note o pienamente comprese al momento della raccolta stessa.

Inoltre, ogni set di dati conterrà probabilmente diversi punti dati o valori sulle persone le cui informazioni personali vengono raccolte.

Il principio dell'accuratezza dei dati richiede che i dati siano completi e aggiornati. Le informazioni dovrebbero essere rappresentative della popolazione target, non includere proxy discriminatori come razza, etnia o religione e risultati intesi solo come correlazioni e non causalità. Il collegamento di dati provenienti da varie fonti può aumentare la probabilità che le decisioni derivanti da tali dati siano basate su informazioni inesatte o sulla base della documentazione storica di un individuo piuttosto che di circostanze attuali.

La distorsione nei grandi set di dati può essere sconosciuta a causa della mancanza di campionamento, o a causa di una scarsa progettazione della ricerca. Se un set di dati contiene una variabile che non è protetta dalla legge ma è discriminatoria, ad esempio una regione geografica che contiene un'alta percentuale di individui con la stessa origine razziale o etnica, le decisioni prese dall'analisi possono essere basate su razza ed etnia. Vi è una crescente preoccupazione che l'uso di tali dati possa costituire una forma di sorveglianza dei dati che opera contro gli interessi legittimi dell'individuo.

Lo sviluppo di algoritmi avanzati ha permesso ai big data di rilevare la presenza di relazioni sempre più complesse tra un numero significativamente elevato di variabili, e questa capacità porta con sé un rischio importantissimo di individuazione dei soggetti ai quali questi dati possono essere riferiti.

La de-identificazione, l'anonimizzazione e la pseudonimizzazione dei dati sono pratiche raccomandate per mitigare il rischio di violazione della privacy in grandi set di dati collegati. Generalmente, si dice che un set di dati viene de-identificato se gli elementi che potrebbero identificare immediatamente una persona o un'organizzazione sono stati rimossi o mascherati.

La legislazione sulla protezione dei dati definisce un trattamento diverso per i dati identificabili e non identificabili; tuttavia, a volte è difficile fare questa distinzione e in particolare con i dati derivati dall'analisi dei big data.

La profilazione dei big data è un tipo di elaborazione automatizzata delle informazioni personali che inserisce le informazioni personali di un individuo in un modello predittivo, che quindi elabora le informazioni in base all'insieme di regole stabilite dal modello per produrre una valutazione o una previsione relativa a uno o più attributi dell'individuo. Ad esempio, può essere utilizzato per valutare o prevedere l'idoneità di un individuo per programmi o servizi. La profilazione non solo elabora le informazioni personali, ma le genera anche, creando un nuovo elemento di informazioni personali che sarà associato all'individuo.

Mentre la profilazione predefinisce gli individui in tipi o categorie in un approccio riduttivo alla comprensione del comportamento umano, la previsione è impostata in un determinato momento e ci si aspetta un certo grado di errore nel risultato.

È importante per le organizzazioni di profilazione dei dati promuovere la trasparenza della logica utilizzata dal modello predittivo e le potenziali conseguenze dei risultati. È auspicabile che le organizzazioni dovrebbero verificare i risultati delle decisioni basate esclusivamente sulla profilazione per garantire che gli individui possano esercitare il loro diritto alla privacy di contestare o rispondere a tali decisioni.

Per sua stessa natura, la profilazione tratta gli individui come oggetti fissi e trasparenti piuttosto che come soggetti dinamici ed emergenti. Una pratica che potrebbe essere consigliata per un'organizzazione che profila le persone è quella di consultare prima le organizzazioni pubbliche e della società civile in merito all'impatto della profilazione proposta e di condurre una valutazione dell'impatto sulla privacy.

Sul punto è opportuno riportare l'orientamento della Comunità Europea che sarà trattato nei paragrafi successivi, in particolar modo con riferimento al Regolamento generale sulla protezione dei dati (GDPR) dell'Unione europea (UE) che verrà preso in considerazione nei paragrafi successivi.

3. Il rischio del “deskilling” professionale in ambito sanitario

Nel considerare qualsiasi tecnologia, i potenziali rischi e danni devono essere valutati rispetto ai potenziali benefici. In questo lavoro si prendono in considerazione questioni tra cui i valori riflessi nei sistemi di IA, la loro spiegabilità, la necessità di concentrarsi sui risultati clinici, il potenziale di bias, il problema della trasferibilità, le preoccupazioni relative ai dati, la responsabilità legale, morale e professionale, l'effetto sull'esperienza del paziente e le possibili spiegazioni per la spinta all'implementazione.

Tra le preoccupazioni più importanti sicuramente rientra quello del deskilling professionale. L'uso di sistemi di supporto decisionale basati sull'apprendimento automatico nelle cliniche può portare a un'eccessiva fiducia nelle capacità dell'automazione, portando al suo uso pesante. A lungo termine, ciò potrebbe ridurre il livello di abilità richiesto dai professionisti per eseguire la maggior parte delle attività di automazione, con gravi conseguenze in caso di guasto del sistema stesso.

Sono necessarie ulteriori ricerche per capire meglio se l'eccessivo affidamento sui sistemi di supporto decisionale basati sull'apprendimento automatico può portare a una sottile perdita di fiducia e influenzare la volontà dei medici di fornire una spiegazione o una diagnosi definitiva.

Il fenomeno definito "dequalificazione professionale" comporta l'eccessivo affidamento degli operatori sanitari ai risultati delle macchine intelligenti, che può portare a una perdita di competenze, soprattutto se protratta nel tempo.

L'eccessivo affidamento sull'uso della tecnologia porta a un graduale declino delle capacità analitiche dei medici e nel tempo delle capacità professionali (un esempio pratico potrebbe essere l'uso di una calcolatrice, che consente di effettuare facilmente calcoli precisi. Velocità di calcolo più veloce, ma allo stesso tempo, riduce la propria capacità di fare calcoli "mentalmente").⁵⁹

L'uso continuativo di tecnologie dotate di "intelligenza artificiale", anche solo per motivi a supporto di scelte diagnostiche e terapeutiche, senza rigorosi protocolli di validazione, porterà nel tempo gli operatori sanitari a "distinguersi" in termini di livelli di competenza acquisiti" e richiedere cure mediche. A riprova, è sufficiente dettagliare i numerosi ricorsi preventivi propugnati ex art. 700 cpc,⁶⁰ di cui il periculum in mora, derivato dai professionisti per provare il danno irreparabile per gravità e per inabilitazione, è appunto la competenza professionale, l'esperienza e la scarsa capacità tecnica conoscenza.

Al contrario, la giurisprudenza ha riconosciuto che quanto più l'incarico è altamente specializzato e costantemente rinnovato, tanto maggiore e più rapida è tale inabilitazione (come impoverimento del proprio bagaglio professionale).

In questo caso, la giurisprudenza tratta addirittura la perdita dell'esperienza come danno "irreparabile".

Il deskilling può anche danneggiare i sistemi sanitari nazionali, che potrebbero "avvizzire" se un sistema dotato di IA restituisce risultati falsi o cessa di funzionare, anche temporaneamente.

⁵⁹ Pasceri, G. *"Intelligenza Artificiale, Algoritmo E Machine Learning: La Responsabilità Del Medico e Dell'amministrazione Sanitaria"* Giuffrè Francis Lefebvre, 2021.

⁶⁰ Si veda Tribunale di Milano, ord. 9.03.2005 in GD, 2005 fasc. 41, 78.

Un'eccessiva fiducia nell'output prodotto da agenti intelligenti può portare inconsapevolmente ad abbandonare la valutazione analitica delle informazioni disponibili da varie fonti.

Questo fenomeno è definito come "legittimità algoritmica" o "affidamento sui risultati prodotti da un algoritmo" che potrebbe persino portare a errori critici che nemmeno i professionisti sono più in grado di rilevare.

In sintesi, l'eccessivo affidamento sull'intelligenza artificiale può determinare diversi fenomeni:

- a) l'eccessiva fiducia si chiama *overconfidence*, che è dovuta al fatto che i professionisti si abituanano a immaginare che un agente intelligente non possa mai sbagliare sulla base di precedenti esperienze positive, e quindi può fare affidamento sulle sue decisioni;
- b) una fiducia incondizionata verso l'intelligenza artificiale può favorire, inoltre, la spersonalizzazione del rapporto medico-paziente;
- c) l'*overconfidence* può determinare nel tempo una perdita di autonomia (*deskilling*) nella capacità di valutare correttamente elementi clinici solitamente riservati agli agenti intelligenti nel caso in cui tali apparecchiature siano temporaneamente non disponibili;
- d) la fiducia incondizionata del medico è favorevole anche alla spersonalizzazione del rapporto medico-paziente; professionisti che non riescono ad abituarsi all'utilizzo del dispositivo per ambiente mentale, tradizione o precedenti esperienze negative, diffidano di esso perché sottovaluta il potere predittivo dell'IA e, quindi, manca di fiducia nell'utilità stessa, non riuscendo ad associare pienamente la propria attività a "smart "tecnologia da integrare;⁶¹

Questi fattori determinano la riduzione della "frontiera delle possibilità" e quindi quella che viene definita "inefficienza" della gestione della sanità pubblica. L'eccessivo affidamento sull'uso della tecnologia può portare a un declino delle capacità analitiche e decisionali dei professionisti. Allo stesso tempo, è stato dimostrato che l'affidamento dei professionisti sugli altri porta a margini di errore più elevati rispetto alle prestazioni medie, riducendo la capacità di fornire spiegazioni e diagnosi chiare.

La letteratura offre già alcuni esempi di questo fenomeno.

Un team di ricercatori della Metropolitan University di Londra ha analizzato 180 mammografie di 50 professionisti e ha scoperto che i medici specialisti più anziani erano del 14,5% meno sensibili nella diagnosi del cancro al seno quando veniva loro presentata l'interpretazione computerizzata di immagini difficili da leggere e quando queste venivano presentate come casi con interpretazioni

⁶¹ Pasceri, G. *"Intelligenza Artificiale, Algoritmo E Machine Learning: La Responsabilità Del Medico e Dell'amministrazione Sanitaria"* Giuffrè Francis Lefebvre, 2021.

più semplici, un sottogruppo di medici meno esperti ha visto solo un aumento dell'1,6% della sensibilità diagnostica grazie all'assistenza informatica.⁶²

Tuttavia è opportuno riportare, in contrasto con la visione distopica, che pone l'accento su ciò che potrebbe andare perso, la visione utopica dell'utilizzo di IA in campo sanitario e in riferimento alle skills professionali che invece sottolinea il potenziale di guadagno.

Secondo questa visione, l'implementazione dell'IA porterebbe a una maggiore qualificazione dei medici, sia direttamente attraverso il miglioramento delle competenze sia indirettamente attraverso la delega di compiti/ruoli.

Un modo in cui l'IA potrebbe direttamente migliorare le competenze dei medici è la personalizzazione della loro istruzione e formazione. In questo caso l'IA ha migliorato le competenze del medico stesso, rimanendo estranea al decisore umano.

L'altra forma di aggiornamento diretto è stata quella dei sistemi di supporto alle decisioni, dove l'aggiunta di un'IA ha aumentato l'accuratezza del processo decisionale umano. A differenza del percorso di potenziamento dell'istruzione e della formazione, in cui l'IA era estrinseca all'uomo, il percorso di potenziamento del supporto decisionale ha creato un cyborg uomo-IA in grado di prendere decisioni migliori rispetto all'uomo o alla macchina da soli:

Il clinico che utilizza il computer per supportare il suo processo decisionale sta beneficiando dell'esperienza di tutti e il ricorso al solo computer non crea quella fiducia di cui credo le persone abbiano bisogno. L'uso dell'intelligenza artificiale per esaminare le radiografie e avvisare i radiologi della patologia, perché non corrisponde al modello, è molto meglio dei radiologi che esaminano tutte le pellicole da soli.

La visione utopica suggerisce anche che l'IA svolga i compiti più adatti all'automazione, così che i medici possano concentrarsi su compiti più soddisfacenti o complessi e/o cercare una formazione specialistica. In questo caso l'IA è estrinseca al medico e gli consente di svolgere compiti più difficili, desiderabili e gratificanti. Ciò potrebbe migliorare l'esperienza dei medici nel loro lavoro, offrire loro nuove opportunità e prevenire il burnout.⁶³

L'upskilling indiretto potrebbe verificarsi anche se l'IA assumesse ruoli clinici, consentendo ai medici di passare a ruoli che ritengono molto più soddisfacenti e complessi e finanziariamente remunerativi.

⁶² Povyakalo, Andrey A., et al. "How to Discriminate between Computer-Aided and Computer-Hindered Decisions" in *Medical Decision Making*, 2013, vol. 33, n. 1, pp. 98-107.

⁶³ Saint James Aquino, Yves, et al. "Utopia versus Dystopia: Professional perspectives on the impact of healthcare Artificial Intelligence on clinical roles and skills" in *International Journal of Medical Informatics*, 2022, p. 104903.

A differenza dell'upskilling diretto, l'upskilling indiretto sembra richiedere almeno una certa dequalificazione. Alcuni interpreti del fenomeno hanno sostenuto che il processo di perdita di competenze e di posti di lavoro per lasciare il posto a nuovi è storicamente parte del lavoro sanitario come pratica in evoluzione.

Questa posizione si basa su almeno due presupposti. In primo luogo, l'introduzione di nuove tecnologie in grado di sostituire alcune mansioni cliniche significa che i medici non le svolgeranno più e che, nel tempo, perderanno le competenze necessarie per svolgerle. La seconda ipotesi è che le competenze perse nel tempo non siano più importanti per lo svolgimento di determinati lavori clinici.

Quando si cercavano le proteine o lo zucchero nelle urine, venivano utilizzati bastoncini di carta che si dovevano immergere e poi confrontarne il colore con una scala. Oggi tutto questo sembra quasi medioevale, perché tutto è diventato quantificato e automatizzato, ma ugualmente i medici continuano a lavorare nei laboratori di biochimica o di biologia clinica.

Sembra che l'obiettivo professionale si sia un po' spostato dal fare cose manuali al supervisionare il processo di produzione.

Alcuni interpreti del fenomeno dell'IA in sanità hanno quindi accolto con favore la dequalificazione, sostenendo che fosse necessaria per far progredire il campo della medicina e del lavoro sanitario.

Tuttavia sembra maggiormente condivisibile, visti i rapidi progressi dell'IA in ambito sanitario, la teoria dispotica ed quindi risulta urgente affrontare il rischio di dequalificazione.

Una soluzione del rischio di deskilling professionale potrebbe passare attraverso l'automatizzazione solo delle mansioni sanitarie più adatte all'automazione oppure sviluppare sistemi di formazione su misura basati sull'IA per migliorare le competenze degli operatori umani; valutare quali mansioni sanitarie, anche se banali, possono essere necessarie per gli operatori umani per continuare a sviluppare o mantenere le loro competenze di ordine superiore, tenendo presente che queste possono variare a seconda delle specialità cliniche; considerare l'automazione delle mansioni di assistenza sanitaria; considerare l'automazione dei compiti laddove ciò consentirebbe ai medici umani di concentrarsi su un lavoro più soddisfacente, con l'avvertenza che i sistemi di IA dovrebbero essere in grado di dimostrare un miglioramento dei risultati per i pazienti e un impatto trascurabile sugli aspetti relazionali e assistenziali dell'assistenza sanitaria.

La ricerca futura potrebbe anche esaminare l'impatto effettivo dell'implementazione dei sistemi di IA sulle competenze cliniche, per stabilire se le pratiche risultanti siano più vicine alla visione distopica o utopica, nonché altri impatti, tra cui gli esiti clinici e l'esperienza del paziente.⁶⁴

4. Il ruolo dell'Europa in riferimento all'utilizzo dell' I.A. – orientamenti per una I.A. affidabile

Solo nel febbraio 2020, la Commissione europea ha pubblicato un Libro bianco che contiene un approccio europeo all'eccellenza e alla fiducia per l'I.A. Contemporaneamente, la stessa ha pubblicato una comunicazione su una strategia europea in materia di dati e una relazione sulla responsabilità dell'Intelligenza artificiale nel settore sanitario, delle implicazioni e sicurezza dell'IA, dell'Internet of Things (IoT) e della robotica⁶⁵.

Il Libro bianco della Commissione, in particolare, sottolinea che *"l'Europa può combinare i suoi punti di forza tecnologici e industriali con un'infrastruttura digitale di alta qualità e un quadro normativo basato sui suoi valori fondamentali per diventare un leader mondiale nell'innovazione nell'economia dei dati e nelle sue applicazioni"* ⁶⁶.

Anche il Consiglio europeo ha tentato di dare una definizione di IA individuandola come "un insieme di scienze, teorie e tecniche il cui scopo è riprodurre da una macchina le capacità cognitive di un essere umano".

Gli sviluppi attuali mirano, ad esempio, a poter affidare a una macchina compiti complessi precedentemente delegati a un essere umano, con specifico riguardo a quei sistemi di elaborazione dati con un elevato grado di "autonomia" nel raggiungimento di obiettivi specifici attraverso sofisticate analisi dell'ambiente in cui operano.

⁶⁴ Saint James Aquino, Yves, et al. *"Utopia versus Dystopia: Professional perspectives on the impact of healthcare Artificial Intelligence on clinical roles and skills"* in International Journal of Medical Informatics, 2022, p. 104903.

⁶⁵ Azzali, E. *"Notes on Recent European Commission Plans for regulation of Artificial Intelligence."* In Management 20, 2020 pp. 20- 22.

⁶⁶ "White Paper on Artificial Intelligence: a European approach to excellence and trust". *European Commission*, commission.europa.eu/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en.

Il Consiglio d'Europa⁶⁷ e l'Unione Europea hanno avviato un'interessante riflessione generale su questi temi, al fine di sviluppare un approccio etico e affidabile all'IA.

Nasce infatti una volontà europea di regolamentare più rigorosamente questi sistemi anche se è necessario affrontare la complessità tecnica dell'IA e di tutte le tecnologie correlate. La proposta di legge sull'IA, infatti, testimonia questo percorso europeo.

Guardando al mondo post-Covid, c'è la speranza istituzionale che l'IA possa essere uno dei principali fattori utili per ricostruire un nuovo concetto di Europa. Tuttavia, allo stesso tempo, è necessario riflettere sul ruolo delle tecnologie nella nostra vita per capire, prima di tutto, il tipo di mondo in cui vorremmo vivere.

La Commissione europea⁶⁸ ha adottato la sua strategia di IA per l'Europa nell'aprile 2018.

In questa comunicazione la Commissione ha lanciato un'iniziativa europea sull'IA che mira, tra l'altro, a garantire un quadro etico e giuridico adeguato, ad esempio creando un'alleanza europea per l'IA e sviluppando orientamenti etici in materia di IA.

La Commissione sottolinea, nella presente comunicazione, la necessità che l'intera Unione europea (UE) dovrebbe adoperarsi per aumentare gli investimenti (pubblici e privati) nell'IA.

Il gruppo di esperti ad alto livello sull'IA (AI Hleg) della Commissione europea, nominato dalla Commissione europea nel giugno 2018 e che è anche il gruppo direttivo dell'Alleanza europea per l'IA, ha pubblicato le linee guida etiche nell'aprile 2019.

Le linee guida promuovono lo slogan "IA affidabile" e contengono sette requisiti chiave che i sistemi di IA devono soddisfare per essere affidabili: agenzia umana e supervisione, robustezza tecnica e sicurezza, privacy e gestione dei dati, trasparenza, diversità, non discriminazione ed equità, benessere ambientale e sociale e responsabilità".⁶⁹

⁶⁷ ad esempio il comitato ad hoc sull'intelligenza artificiale (CAHAI), "Prima relazione inter Consiglio d'Europa, "Raccomandazione del Comitato dei Ministri agli Stati membri sull'impatto dei sistemi algoritmici sui diritti umani"

⁶⁸ Si veda: "Commissione europea. Comunicazione della Commissione al Parlamento europeo, al Consiglio europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle regioni. L'intelligenza artificiale per l'Europa", consultabile in <https://eur-lex.europa.eu/legal-content/IT/TXT/?uri=CELEX%3A52018DC0237>.

⁶⁹ "Ethics Guidelines for Trustworthy Ai." Shaping Europe's Digital Future, consultabile <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>.

Ai fini dei suoi risultati, il gruppo di esperti ad alto livello sull'IA ha anche pubblicato un documento sulla definizione di IA⁷⁰. Inoltre, nel giugno 2019, l'IA HLEG ha pubblicato un altro risultato che fornisce "Raccomandazioni politiche e di investimento per un'IA affidabile".

La Commissione europea incoraggia tutti gli Stati membri dell'UE a sviluppare una strategia nazionale in materia di IA, e diversi Stati ne hanno già pubblicata una, come il Regno Unito (UK) e la Germania⁷¹.

Nel dicembre 2018 la Commissione europea ha inoltre concordato un piano coordinato sull'IA con gli Stati membri dell'UE, la Norvegia e la Svizzera per promuovere lo sviluppo e l'utilizzo dell'IA in Europa. L'obiettivo generale della collaborazione è garantire che l'Europa diventi la regione leader mondiale per lo sviluppo e l'applicazione di un'IA all'avanguardia, etica e sicura⁷².

Il gruppo di esperti ad alto livello sull'intelligenza artificiale crede che l'IA abbia il potenziale per trasformare in modo significativo la società. L'IA non è un fine in sé, ma un mezzo promettente per aumentare la prosperità umana, migliorando così il benessere individuale e sociale e il bene comune, oltre a portare progresso.

In particolare, i sistemi di IA possono contribuire a facilitare gli obiettivi di Sviluppo Sostenibile delle Nazioni Unite, come la promozione dell'equilibrio di genere e la lotta al cambiamento climatico, razionalizzando l'uso delle risorse naturali, migliorare la salute, la mobilità e i processi di produzione e sostenere le modalità di monitoraggio dei progressi in materia di sostenibilità e coesione sociale.

A tal fine, i sistemi di IA devono essere incentrati sull'uomo, basandosi sull'impegno a utilizzarli al servizio dell'umanità e del bene comune con l'obiettivo di migliorare il benessere e la libertà dell'uomo.

Pur offrendo grandi opportunità, i sistemi di IA comportano anche alcuni rischi che devono essere gestiti in modo appropriato e proporzionato. La Commissione europea si pone come obiettivo quello di massimizzare i vantaggi dei sistemi di IA e allo stesso tempo di prevenire e minimizzare i loro rischi.

⁷⁰ Floridi, Luciano. "Establishing the Rules for Building Trustworthy AI" in Electronic Journal, 2019, pp. 261–262, <https://doi.org/10.1038/s42256-019-0055-y>

⁷¹ "ETIM 2020 – Künstliche Intelligenz und MicroLearning". *RöFo - Fortschritte auf dem Gebiet der Röntgenstrahlen und der bildgebenden Verfahren*, vol. 192, n. 02, gen. 2020, p. 137

⁷² "Member States and Commission to work together to boost artificial intelligence “made in Europe”". *Shaping Europe’s digital future*, digital-strategy.ec.europa.eu/en/news/member-states-and-commission-work-together-boost-artificial-intelligence-made-europe.

Le linee guida sostengono un approccio responsabile allo sviluppo dell'IA, che dovrebbe essere legale, rispettando tutte le leggi e i regolamenti applicabili; etico, rispettando i principi e i valori etici e robusto, sia dal punto di vista tecnico che da quello sociale.

- Legale: i sistemi di IA non operano in un mondo senza leggi. Una serie di norme giuridicamente vincolanti a livello europeo, nazionale e internazionale si applicano già o sono rilevanti per lo sviluppo, la diffusione e l'utilizzo dei sistemi di IA oggi.

Le fonti giuridiche includono: il diritto primario dell'UE (i Trattati dell'Unione europea e la Carta dei diritti fondamentali), il diritto derivato dell'UE (come il Regolamento generale sulla protezione dei dati, la Direttiva sulla responsabilità dei prodotti, il Regolamento sulla libera circolazione dei dati non personali, le direttive antidiscriminazione, il diritto dei consumatori e le direttive sulla sicurezza e la salute sul lavoro), i trattati delle Nazioni Unite sui diritti umani e le convenzioni del Consiglio d'Europa (come la Convenzione europea dei diritti dell'uomo) e numerose leggi degli Stati membri dell'UE.

Oltre alle norme applicabili orizzontalmente, esistono diverse norme specifiche per ogni settore che si applicano a particolari applicazioni dell'IA (come ad esempio il regolamento sui dispositivi medici nel settore sanitario).

La legge prevede obblighi sia positivi che negativi, il che significa che non deve essere interpretata solo con riferimento a ciò che non può essere fatto, ma anche con riferimento a ciò che dovrebbe essere fatto e a ciò che può essere fatto.

La legge non solo proibisce alcune azioni, ma ne consente altre. A questo proposito, si può notare che la Carta dell'UE contiene articoli sulla "libertà d'impresa" e sulla "libertà delle arti e delle scienze", oltre ad articoli che affrontano aree che ci sono più familiari quando si cerca di garantire l'affidabilità dell'IA, come ad esempio la protezione dei dati e la non discriminazione. Le Linee guida non trattano esplicitamente la prima componente dell'IA affidabile (IA lecita), ma mirano invece a offrire indicazioni per promuovere e garantire la seconda e la terza componente (IA etica e robusta).

Sebbene queste ultime due componenti siano in una certa misura già riflesse nelle leggi esistenti, la loro piena realizzazione può andare oltre gli obblighi legali esistenti. Nulla di quanto contenuto nelle linee guida dovrebbe essere interpretato come una consulenza o una guida legale su come raggiungere la conformità con le norme e i requisiti legali vigenti.

Nulla di quanto contenuto nel presente documento creerà diritti legali né imporrà obblighi legali nei confronti di terzi.

Si ricorda tuttavia che è dovere di ogni persona fisica o giuridica rispettare le leggi, siano esse applicabili oggi o adottate in futuro in base allo sviluppo dell'IA.

Le Linee guida analizzate partono dal presupposto che tutti i diritti e gli obblighi legali che si applicano ai processi e alle attività coinvolte nello sviluppo, nell'implementazione e nell'utilizzo dei sistemi di IA, rimangono obbligatori e devono essere debitamente osservati.

- Etico: il raggiungimento di un'IA affidabile non richiede solo il rispetto della legge, che è solo una delle sue tre componenti. Le leggi non sono sempre al passo con gli sviluppi tecnologici, a volte non sono al passo con le norme etiche o semplicemente non sono adatte ad affrontare determinate questioni. Per essere affidabili, i sistemi di IA dovrebbero quindi essere anche etici, garantendo l'allineamento con le norme etiche.
- Tecnico e Sociale: anche se viene garantito uno scopo etico, gli individui e la società devono avere la certezza che i sistemi di IA non causeranno alcun danno involontario. Tali sistemi dovrebbero funzionare in modo sicuro, protetto e affidabile e dovrebbero essere previste misure di salvaguardia per prevenire eventuali impatti negativi non intenzionali. È quindi importante garantire la solidità dei sistemi di IA. Ciò è necessario sia dal punto di vista tecnico (garantendo la robustezza tecnica del sistema come appropriato in un determinato contesto, come il dominio applicativo o la fase del ciclo di vita), sia dal punto di vista sociale (tenendo in debita considerazione il contesto e l'ambiente in cui il sistema opera). L'IA etica e quella robusta sono quindi strettamente interconnesse e si completano a vicenda.

Poiché l'IA diventerà sempre più importante e pervasiva, deve funzionare in modo affidabile, in modi che tutti possano ritenere utili per l'umanità e per l'intero ambiente. L'alternativa è che l'IA possa essere usata in modo improprio, eccessivo o insufficiente.

L'incertezza etica genera sia l'assunzione di rischi sconsiderati che l'eccessiva cautela. Ecco perché le linee guida sono così importanti. Rappresentano un buon passo nella giusta direzione di un quadro di riferimento chiaro, condiviso e socialmente preferibile per l'IA etica.

Le linee guida sono state sia elogiate e accolte con favore che criticate⁷³ per la loro debolezza, in quanto fanno parte di una mera strategia di autoregolamentazione, che non viene applicata legalmente, e per la loro scarsa utilità, in quanto sono troppo generiche e si aggiungono a tante altre iniziative che finora hanno avuto scarso impatto.

Queste e altre critiche possono essere confutate.

In primo luogo, le linee guida contengono principi e chiarimenti solidi, in termini di aspettative sociali, e coerenti con lo stato attuale del dibattito sull'etica dell'IA. Naturalmente, sono necessarie

⁷³ Der Tagesspiegel “*EU guidelines: Ethics Washing Made in Europe*” Aktuelle News: Nachrichten aus Berlin und der Welt. 2019, consultabile al <https://www.tagesspiegel.de/politik/ethics-washing-made-in-europe-5937028.html>.

sia la legge che l'etica dell'IA. Le linee guida presuppongono e sono allineate con la legislazione dell'UE.

L'UE è in prima linea nel dibattito internazionale sull'IA, anche grazie al Regolamento generale sulla protezione dei dati (GDPR). L'etica può contribuire alla definizione di una nuova legislazione (ad esempio, sui sistemi di riconoscimento facciale) o fungere da guida in sua assenza. A volte, l'etica è necessaria per interpretare la legislazione esistente (ad esempio, il GDPR).

Altre volte, l'etica può raccomandare di non fare qualcosa che la legislazione non vieta (per esempio, lasciare una decisione medica interamente a un algoritmo senza supervisione o spiegazione), o raccomandare di fare qualcosa che la legislazione non richiede (per esempio, progettare un algoritmo che minimizzi l'impatto ambientale del riscaldamento domestico). In tutti questi casi, la conformità alla legge è necessaria ma insufficiente e, come riconoscono le linee guida, deve essere integrata da un approccio di "etica morbida" successivo alla conformità⁷⁴, perché la legge fornisce le regole del gioco, ma non indica come giocare bene secondo le regole.

In secondo luogo, è condivisibile pensare che le linee guida non siano molto originali o innovative, ma questo è comunque giustificabile dalle discussioni sull'argomento dell'etica dell'IA che vanno avanti da quasi mezzo secolo⁷⁵.

Attualmente, infatti, esistono più di 70 quadri ed elenchi di principi sull'etica dell'IA⁷⁶.

Le linee guida sono la cosa più vicina a uno standard completo e autorevole disponibile nell'UE, offrendo un chiaro quadro di riferimento e un vocabolario concettuale comune.

Sono state concepite per stabilire un punto di riferimento per ciò che può o non può essere considerato, d'ora in poi, un'IA affidabile.

Una parte delle linee guida si occupa dei diritti fondamentali quali diritti morali e diritti giuridici.

⁷⁴ Floridi, Luciano, e Mariarosaria Taddeo. "What is data ethics?" in *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 2016, vol. 374, n. 2083, p. 20160360.

⁷⁵ Wiener, N. "Some Moral and Technical Consequences of Automation" in *Science*, 1960 vol. 131, n. 3410, pp. 1355-58

⁷⁶ "AI Ethics Guidelines Global Inventory - AlgorithmWatch" AlgorithmWatch, algorithmwatch.org/en/project/ai-ethics-guidelines-global-inventory. Ultimo accesso il 20 feb. 2023.

Come già evidenziato la Commissione confida in un approccio all'IA basato sui diritti fondamentali sanciti dai trattati dell'UE⁷⁷, dalla Carta dei diritti fondamentali dell'UE e dal diritto internazionale in materia di diritti umani⁷⁸.

Il rispetto dei diritti fondamentali nel quadro della democrazia e dello stato di diritto costituisce la base più ispiratrice per identificare principi e valori etici astratti che possono operare nei sistemi di IA.

I trattati dell'UE e la Carta dell'UE stabiliscono una serie di diritti fondamentali che gli Stati membri e le istituzioni dell'UE sono giuridicamente obbligati a rispettare nell'attuazione del diritto dell'UE. Questi diritti sono descritti nella Carta dei diritti fondamentali dell'Unione europea e riguardano la dignità, la libertà, l'uguaglianza, la solidarietà, i diritti civili e la giustizia. La base per unificare questi diritti può essere intesa come radicata nel rispetto della dignità della persona umana, riflettendo così quello che definiamo un "approccio antropocentrico" in cui gli esseri umani godono di un primato unico e inalienabile dello status morale.

Sebbene i diritti sanciti dalla Carta dei diritti fondamentali dell'UE siano giuridicamente vincolanti⁷⁹, è importante riconoscere che i diritti fondamentali non sempre includono tutele giuridiche complete.

Ad esempio, per quanto riguarda la Carta dei diritti fondamentali dell'UE, è importante sottolineare che il suo campo di applicazione è limitato alle questioni di diritto dell'UE. Il diritto internazionale dei diritti umani, in particolare la Convenzione europea dei diritti dell'uomo, è giuridicamente vincolante per gli Stati membri dell'UE, comprese le questioni al di fuori dell'ambito di applicazione del diritto dell'UE.

Allo stesso tempo, va sottolineato che i diritti fondamentali sono conferiti anche agli individui in conformità con il loro status morale di esseri umani (in una certa misura), indipendentemente dalla loro validità giuridica. Concepiuti come diritti giuridicamente vincolanti, i diritti fondamentali fanno quindi parte della prima componente di un'IA affidabile (la legittimità dell'IA), che garantisce il rispetto della legge. Sono intesi come i diritti di ogni essere umano, radicati nello stato morale

⁷⁷ L'UE si fonda su un impegno costituzionale a tutelare i diritti fondamentali e indivisibili degli esseri umani, a garantire il rispetto dello Stato di diritto, a promuovere la libertà democratica e il bene comune. Questi diritti sono sanciti negli articoli 2 e 3 del trattato sull'Unione europea e nella Carta dei diritti fondamentali dell'UE.

⁷⁸ Tali impegni sono ulteriormente specificati in altri strumenti giuridici, quali ad esempio la Carta sociale europea del Consiglio d'Europa o in atti legislativi specifici come il regolamento generale dell'UE sulla protezione dei dati.

⁷⁹ A norma dell'articolo 51, le disposizioni della Carta si applicano alle istituzioni dell'Unione come pure agli Stati membri nell'attuazione del diritto dell'Unione.

intrinseco degli esseri umani, e sono anche la base per la seconda componente dell'intelligenza artificiale affidabile (etica dell'intelligenza artificiale), che coinvolge norme etiche che non sono necessariamente legalmente vincolante, ma fondamentale per garantire l'affidabilità.

5. Il GDPR e il diritto alla spiegazione

Il Regolamento generale sulla protezione dei dati (GDPR) dell'Unione europea (UE), pienamente applicabile da maggio 2018, sostituisce la direttiva sulla protezione dei dati del 1995 e rafforza e armonizza la protezione dei dati personali per i cittadini dell'UE.

Questa legislazione ha una portata globale e ha effettivamente influenzato i cambiamenti legislativi nella protezione della privacy.

Il GDPR ha ampliato i principi di protezione dei dati per richiedere alle organizzazioni di dimostrare responsabilità nella raccolta, nell'uso e nella divulgazione delle informazioni personali.

L'importanza emergente della responsabilità è in risposta diretta alle implicazioni del trattamento dei dati personali in un mondo di big data. Più specificamente, il GDPR richiede che sia completata una valutazione d'impatto sulla protezione dei dati per le iniziative che comportano "una valutazione sistematica ed estesa degli aspetti personali relativi alle persone fisiche che si basa sul trattamento automatizzato, compresa la profilazione, e su cui si basano decisioni che producono effetti giuridici riguardanti la persona fisica o influenzano in modo analogo significativo la stessa".⁸⁰

Altre disposizioni del regolamento includono la protezione dei dati fin dalla progettazione e per impostazione predefinita, ad esempio la privacy by design e la certificazione oppure l'istituzione di meccanismi di certificazione e sigilli e marchi di protezione dei dati per le organizzazioni per fornire un rapido accesso pubblico al livello di protezione dei dati di prodotti e servizi pertinenti.

Un'opinione prevalente è che qualsiasi potenziale danno derivante dall'analisi dei big data derivi dal modo in cui i dati vengono utilizzati, non necessariamente dal modo in cui i dati sono stati raccolti. Il principio di responsabilità del GDPR ora si concentra sull'uso delle informazioni personali

⁸⁰*Regolamento Generale Sulla Protezione Dei Dati. Home - Garante Privacy*, si veda www.garanteprivacy.it/documents/10160/0/Regolamento+UE+2016+679.+Arricchito+con+riferimenti+ai+Considerando+Aggiornato+alle+rettifiche+pubblicate+sulla+Gazzetta+Ufficiale++dell%27Unione+europea+127+del+23+maggio+2018.

attraverso meccanismi come il controllo della progettazione tecnica degli algoritmi, la verificabilità del processo di analisi e l'applicazione di regolamenti software-defined.

La responsabilità è stata sostenuta rispetto alla trasparenza che fino ad oggi è nota per avere molte limitazioni nella protezione del diritto alla privacy di un individuo.

In un recente rapporto, l'Information Commissioner per il Regno Unito⁸¹ ha proposto le seguenti sei raccomandazioni per le organizzazioni che conducono analisi dei big data:

1. Valutare attentamente se l'analisi dei big data richiede il trattamento dei dati personali e utilizzare tecniche appropriate per rendere anonimi i dati personali nei set di dati prima dell'analisi.
2. Essere trasparenti sul trattamento dei dati personali utilizzando una combinazione di approcci per fornire informative sulla privacy significative nelle fasi appropriate di un progetto di big data. Ciò può includere l'uso di icone, notifiche just-in-time e informative sulla privacy a più livelli.
3. Integrare un quadro di valutazione dell'impatto sulla privacy nelle attività di trattamento dei big data per aiutare a identificare i rischi per la privacy e valutare la necessità e la proporzionalità di un determinato progetto.

La valutazione dell'impatto sulla privacy dovrebbe coinvolgere il contributo di tutte le parti interessate, compresi gli analisti dei dati, i responsabili della conformità, i membri del consiglio di amministrazione e il pubblico.

4. Adottare un approccio privacy by design nello sviluppo e nell'applicazione dell'analisi dei big data. Ciò dovrebbe includere l'implementazione di misure tecniche e organizzative per affrontare la sicurezza dei dati, la minimizzazione dei dati e la segregazione dei dati.
5. Sviluppare principi etici per contribuire a rafforzare i principi fondamentali di protezione dei dati. Le organizzazioni più grandi dovrebbero creare comitati etici per aiutare a esaminare i progetti e valutare problemi complessi derivanti dall'analisi dei big data.
6. Implementare tecniche innovative per sviluppare algoritmi di apprendimento automatico verificabili. Dovrebbero essere effettuati audit interni ed esterni al fine di spiegare la logica alla base delle decisioni algoritmiche e verificare la presenza di pregiudizi, discriminazioni ed errori.⁸²

⁸¹De Vinco A., "L' utilizzo degli M.M.S. e la protezione dei dati personali nel Regno Unito" AmbienteDiritto, 2006, si veda www.ambientediritto.it/dottrina/Dottrina_2006/utilizzo_M.M.S_devinco.htm.

⁸² Pattelli, Ivano. "Privacy, Sicurezza Ed Etica Nell'utilizzo Dei Big Data e Dell'intelligenza Artificiale." ICT Security Magazine, 2022, si veda <https://www.ictsecuritymagazine.com/articoli/privacy-sicurezza-ed-etica-nellutilizzo-dei-big-data-e-dellintelligenza-artificiale/>

I regimi di protezione dei dati personali, come il GDPR, sono strumenti per la governance dei flussi e del trattamento dei dati e rimangono preziosi per la protezione dei dati personali in linea con il classico trattamento dei dati. Tuttavia, potrebbero essere inadeguati per affrontare le sfide senza precedenti sollevate dai big data. In particolare, la frequente incompatibilità tra big data e principi di privacy. Gli scopi dell'analisi dei big data basata su algoritmi sono spesso quelli di scoprire modelli altrimenti invisibili nei dati, piuttosto che applicare approfondimenti precedenti, testare ipotesi o sviluppare spiegazioni. Si aggiunga a ciò le complessità tecniche dell'apprendimento automatico e dell'intelligenza artificiale, e l'effetto può essere l'allontanamento delle autorità e delle imprese di vigilanza dal significato del diritto alla protezione dei dati.

L'etica consente questo ritorno allo spirito della legge e offre altri spunti per condurre un'analisi di una nuova società digitale, come il suo ethos collettivo, le sue rivendicazioni di giustizia sociale, democrazia e libertà personale.

Nel mondo dei big data, è di fondamentale importanza che esistano leggi sulla protezione dei dati che proteggano adeguatamente la privacy degli individui, in particolare dei pazienti.

Come già sottolineato il Regolamento Generale sulla protezione dei dati introdotto una nuova era della legge sulla protezione dei dati nell'UE.

Per quanto riguarda il suo ambito di applicazione, il GDPR si rivolge al "trattamento dei dati personali nel contesto delle attività di uno stabilimento di un responsabile del trattamento o di un responsabile del trattamento" nell'UE, indipendentemente dal fatto che il trattamento avvenga in un paese dell'UE o non UE, come negli Stati Uniti (artt. 2, 3 del GDPR).

Il GDPR si applica anche quando un responsabile del trattamento tratta dati personali ed è stabilito in un paese terzo, ma "in un luogo in cui si applica il diritto dello Stato membro in virtù del diritto internazionale pubblico" (articolo 3, paragrafo 3, del GDPR).

Per questo motivo, il GDPR potrebbe avere implicazioni anche per le aziende statunitensi. Ad esempio, il Regolamento si applica nei casi in cui il responsabile del trattamento o il responsabile del trattamento sia stabilito in un paese extra UE e tratti "dati personali di interessati che si trovano nell'Unione" per "l'offerta di beni o servizi" (ad esempio, giornali e siti Web affiliati gratuitamente o a pagamento) a tale interessato nell'UE o per il "monitoraggio" del comportamento degli interessati (articolo 3, paragrafo 2, del GDPR).

A questo punto sembra opportuno riportare alcune definizioni essenziali contenute nel citato regolamento per delineare la normativa sulla privacy e per capire come questo incida nel campo medico.

Il termine "dati personali" è definito come "qualsiasi informazione relativa a una persona fisica identificata o identificabile ("interessato")" (articolo 4, paragrafo 1, del GDPR).

Il GDPR definisce "trattamento" come "qualsiasi operazione o insieme di operazioni, compiute con o senza l'ausilio di processi automatizzati e applicate a dati personali o insiemi di dati personali", compresa la raccolta, la strutturazione, la conservazione o l'utilizzo (articolo 4, paragrafo 2, del GDPR). Mentre per "responsabile del trattamento" si intende "la persona fisica o giuridica, l'autorità pubblica, il servizio o altro organismo che, singolarmente o insieme ad altri, determina le finalità e i mezzi del trattamento di dati personali", per "incaricato del trattamento" si intende "la persona fisica o giuridica, l'autorità pubblica, il servizio o altro organismo che tratta dati personali per conto del responsabile del trattamento" (art. 4, n. 7 del GDPR).

Nel contesto sanitario, la definizione di "dati relativi alla salute" è dettata dall'articolo 4, paragrafo 15, del GDPR. Quest'ultimi vengono individuati come "dati personali relativi alla salute fisica o mentale di una persona fisica, compresa la fornitura di servizi sanitari, che rivelano informazioni sul suo stato di salute".

Il GDPR dell'UE è quindi molto più ampio nel suo ambito rispetto all'HIPAA degli Stati Uniti, che copre solo specifiche informazioni sanitarie, come sarà analizzato nei paragrafi successivi.

Ai sensi dell'articolo 9, paragrafo 1, del GDPR, è vietato il trattamento di categorie speciali di dati personali come dati genetici⁸³, dati biometrici⁸⁴ e dati relativi alla salute.

Tuttavia, l'articolo 9, paragrafo 2, del GDPR contiene un elenco di eccezioni. Ad esempio, il divieto di cui all'articolo 9, paragrafo 1, del GDPR di norma non si applica nei casi in cui "l'interessato ha dato il consenso esplicito (...) per una o più finalità determinate" o quando il "trattamento è necessario per motivi di interesse pubblico nel settore della sanità pubblica" o "a fini di archiviazione nel pubblico interesse, a fini di ricerca scientifica o storica o a fini statistici"⁸⁵.

Gli Stati membri dell'UE possono anche decidere di introdurre o mantenere ulteriori requisiti, comprese limitazioni, ma solo "per quanto riguarda il trattamento di dati genetici, dati biometrici o dati relativi alla salute"⁸⁶.

Il mancato rispetto delle presenti condizioni del GDPR potrebbe comportare sanzioni amministrative⁸⁷.

Le prime multe in ambito sanitario sono già state imposte ai sensi del GDPR. Ad esempio, a un ospedale in Portogallo sono stati addebitati 400 mila euro per due violazioni del GDPR: in primo luogo, 300 mila euro per il permesso di "accesso indiscriminato a un insieme di dati da parte di

⁸³ articolo 4, paragrafo 13, del GDPR

⁸⁴ articolo 4, paragrafo 14, del GDPR

⁸⁵ art. 9(2)(a), (i) e (j) del GDPR

⁸⁶ articolo 9, paragrafo 4, del GDPR

⁸⁷ articolo 83, paragrafo 5, del GDPR

professionisti, che dovrebbero potervi accedere solo in casi specifici"; e in secondo luogo, 100 mila euro per l'incapacità di "garantire la riservatezza, l'integrità, la disponibilità e la resilienza permanente dei sistemi e dei servizi di trattamento" ⁸⁸.

Il GDPR contiene anche disposizioni che sono particolarmente rilevanti per la medicina infusa di AI.

Infatti, come già ampiamente individuate nei paragrafi precedenti, esistono diverse applicazioni di I.A.: dalla diagnosi precoce, al supporto nelle operazioni; dalla gestione di database enormi o condivisi, allo studio di terapie personalizzate. Tuttavia, nonostante l'interesse dei mass media per questo tema, la capacità di autoapprendimento e l'autonomia sono ancora nelle loro fasi iniziali.

Ci sono già applicazioni sanitarie AI in Europa. Ad esempio, Ada è un'app sanitaria AI che valuta i sintomi di un individuo e fornisce indicazioni (ad esempio, suggerire all'utente una visita da un medico o di cercare cure di emergenza).

Nell'agosto 2018, i ricercatori di DeepMind e Moorfields Eye Hospital di Londra, Regno Unito, hanno pubblicato su Nature Medicine i risultati dello studio di un sistema di intelligenza artificiale in grado di leggere le scansioni oculari e formulare raccomandazioni di riferimento, comprendente oltre 50 diagnosi comuni. ⁸⁹

Un altro esempio è l'Ultromics. Il team dell'Università di Oxford si è dedicato a ridurre le diagnosi errate e consentire la prevenzione precoce delle malattie cardiovascolari. EchoGo Pro di Ultromics, ⁹⁰ ad esempio, è un sistema di intelligenza artificiale basato sui risultati con marcatura CE in Europa che prevede la malattia coronarica in una fase precoce.

Corti è un software sviluppato da una società danese che sfrutta il ML per aiutare i dispatcher di emergenza a prendere decisioni. Questo strumento è in grado di rilevare gli arresti cardiaci extraospedalieri (cioè quelli che si verificano in pubblico o a casa) durante le chiamate di emergenza

⁸⁸Lusa. "Hospital Do Barreiro Contesta Judicialmente Coima De 400 Mil Euros De Comissão De Dados." Público, 2018, si veda <https://www.publico.pt/2018/10/22/sociedade/noticia/hospital-barreiro-contesta-judicialmente-coima-400-mil-euros-comissao-dados-1848479>

⁸⁹ Gadmin. "Intelligenza Artificiale in Medicina, L'esempio Di Deepmind in Oculistica." Galeno, 2020, <https://cassagaleno.eu/intelligenza-artificiale-in-medicina-lesempio-di-deepmind-in-oculistica/>

⁹⁰ "Artificial Intelligence Improves Coronary Artery Disease Detection, New Study Finds." Home, Retrieved, 2023, si veda <https://www.bsecho.org/Public/Public/News/Articles/2021/2021-12/202112-Ultromics.aspx>

in modo più rapido e accurato rispetto agli esseri umani ascoltando le chiamate e analizzando i sintomi, il tono della voce, i respiratori e altri metadati in tempo reale.⁹¹

In relazione alle applicazioni mediche, al fine di individuare le direttive e i criteri di valutazione, è necessario verificare quale sia il margine di apprezzamento che l'ordinamento giuridico riserva all'autonomia decisionale.

La liceità e i confini dell'autonomia sono, infatti, le variabili contemplate dal sistema europeo come punti di riferimento. In quest'ottica, gli studiosi hanno esaminato il rapporto tra protezione dei dati e processo decisionale autonomo.

Innanzitutto, sembra che tutti i sistemi diagnostici basati sull'intelligenza artificiale siano conformi al GDPR quando formulano ipotesi sullo stato clinico del paziente attraverso l'analisi di imaging intelligente e i loro risultati vengono convalidati da un medico, prima di raggiungere il paziente.

In questo caso, l'IA agisce allo stesso modo di qualsiasi altro software nella piena disponibilità dell'operatore; pertanto, potrebbe essere qualificato come un normale dispositivo medico.

Infatti, non è il potere di calcolo che solleva dubbi sulla liceità (e appropriatezza) dell'uso dell'IA, ma l'indipendenza dagli esseri umani.

Pertanto, le diagnosi e le potenziali terapie derivanti dall'elaborazione dell'IA possono essere considerate linee guida mediche: la decisione algoritmica, filtrata dall'operatore sanitario, potrebbe essere qualificata come protocollo medico. In questo scenario, la presenza di un filtro umano sostanziale porta ad escludere l'applicabilità dell'articolo 22 GDPR.

Tuttavia, il diverso caso di una più ampia indipendenza dell'IA, richiede un focus specifico. Un algoritmo autonomo potrebbe legittimamente adottare una decisione significativa in ambito sanitario? Tutto dipende da come viene concepito il lavoro medico. Se si qualifica il medico come un semplice esecutore sillogistico di linee guida e protocolli, la risposta può essere affermativa. Da questo punto di vista, un algoritmo potrebbe essere ancora più competente ed efficace di un essere umano. Tuttavia, si ritiene che l'azione medica sia caratterizzata da una maggiore complessità valutativa e decisionale. E poiché l'ordinamento giuridico presenta una certa diffidenza verso tutto ciò che sfugge drasticamente al controllo umano, considerata la necessità di garantire un certo standard di sicurezza e affidabilità, sembra necessario ricercare l'approccio da seguire, tenendo conto anche dell'assenza di soluzioni normative specifiche.

⁹¹ “Corti, *Quando L'intelligenza Artificiale Salva Vite Umane.*” *01net*, 2019, si veda <https://www.01net.it/corti-intelligenza-artificiale-salva-vite-umane/>

In riferimento alla tutela della privacy nell'ambito sanitario ed in particolare quando sono raccolti dati personali, i responsabili del trattamento devono generalmente fornire agli interessati informazioni sull'esistenza di un processo decisionale automatizzato, compresa la profilazione, di cui all'articolo 22, paragrafi 1 e 4, e, almeno in tali casi, informazioni significative sulla logica utilizzata, nonché sull'importanza e le conseguenze previste di tale trattamento per l'interessato. Inoltre, gli interessati hanno il diritto di accedere ai dati personali che li riguardano e alle informazioni sull'esistenza di un processo decisionale automatizzato, compresa la profilazione, (...) e (...) informazioni significative sulla logica utilizzata, nonché sul significato e le conseguenze previste di tale trattamento per l'interessato.

Per "processo decisionale automatizzato" si intende una decisione presa – senza alcun coinvolgimento umano – esclusivamente con mezzi automatizzati.

Il termine "profilazione" è definito all'articolo 4, paragrafo 4, del GDPR come "qualsiasi forma di trattamento automatizzato di dati personali consistente nell'uso di dati personali per valutare determinati aspetti personali relativi a una persona fisica, in particolare per analizzare o prevedere aspetti riguardanti le prestazioni lavorative, la situazione economica, la salute, le preferenze personali di tale persona fisica, interessi, affidabilità, comportamento, posizione o movimenti." Pertanto, il termine "profilazione" è un sottoinsieme del termine "trattamento" con due requisiti aggiuntivi, vale a dire che il trattamento deve essere (1) automatizzato e (2) a fini di valutazione.

Ai sensi dell'articolo 22, paragrafo 1, del GDPR, gli interessati hanno inoltre "il diritto di non essere sottoposti a una decisione basata unicamente sul trattamento automatizzato, compresa la profilazione, che produca effetti giuridici che lo riguardano o incida in modo analogo significativamente su di lui o su di lei".⁹²

L'articolo 22, paragrafo 2, del GDPR elenca alcune eccezioni all'articolo 22, paragrafo 1, del GDPR, ma tali eccezioni non si applicano generalmente quando le decisioni si basano su dati genetici e biometrici nonché su dati relativi alla salute.⁹³

Il paragrafo 4 del citato articolo, del GDPR vieta le decisioni automatizzate basate su categorie particolari di dati personali «a meno che non si applichi l'articolo 9, paragrafo 2, lettere a) o g), e

⁹² *Regolamento Generale Sulla Protezione Dei Dati.*
<https://www.garanteprivacy.it/documents/10160/0/Regolamento+UE+2016+679.+Arricchito+con+riferimenti+ai+Considerando+Aggiornato+alle+rettifiche+pubblicate+sulla+Gazzetta+Ufficiale++dell%2527Unione+europea+127+del+23+maggio+2018>

⁹³ articolo 22, paragrafo 4, del GDPR

non siano in atto misure adeguate per salvaguardare i diritti, le libertà e gli interessi legittimi dell'interessato».

Il prodotto dell'elaborazione algoritmica deve raggiungere un certo livello di «significatività che deve essere simile a una decisione che produce un effetto giuridico». A causa della difficile identificazione di un effetto derivante dal trattamento di dati che, interessando la persona, non è giuridicamente rilevante una decisione in grado di influenzare direttamente la salute o il benessere delle persone (ma anche il relativo diritto) può essere qualificata come decisione.

Un esempio concreto potrebbe essere un protocollo terapeutico, monitorato tramite app, che deve aiutare il paziente nel contesto delle procedure di follow-up dopo l'impianto di un pacemaker.

È molto controverso, tuttavia, se il GDPR conceda effettivamente un "diritto alla spiegazione" e cosa significhi tale diritto. L'art. 71 del GDPR menziona esplicitamente il diritto di ottenere una spiegazione della decisione raggiunta tramite macchina.

Alcuni studiosi dubitano dell'esistenza giuridica e della fattibilità di tale diritto alla spiegazione di specifiche decisioni automatizzate, tra l'altro perché l'art. 71 del GDPR non è giuridicamente vincolante e il diritto alla spiegazione non è imposto dai requisiti giuridicamente vincolanti di cui all'articolo 22, paragrafo 3, del GDPR.⁹⁴

Pertanto, secondo questo punto di vista, non esiste fin dall'inizio alcun diritto giuridicamente vincolante dell'interessato di ricevere informazioni sul processo decisionale interno degli algoritmi⁹⁵, e quindi di aprire le "scatole nere" delle applicazioni di IA sanitaria. Tuttavia, se non esiste un diritto giuridicamente vincolante alla spiegazione di specifiche decisioni automatizzate, gli articoli 13, paragrafo 2, lettera f), 14, paragrafo 2, lettera g), e 15, paragrafo 1, lettera h), del GDPR autorizzano almeno gli interessati a ottenere "informazioni significative sulla logica utilizzata, nonché sul significato e le conseguenze previste" dei sistemi decisionali automatizzati. Queste informazioni includono lo scopo di un sistema decisionale automatizzato, il funzionamento del sistema in generale, l'impatto previsto e altre funzionalità del sistema come gli alberi decisionali e le strutture di classificazione⁹⁶.

⁹⁴ Wachter, Sandra, et al. "Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation". SSRN Electronic Journal, 2016.

⁹⁵ Wachter, Sandra, et al. "Counterfactual Explanations Without Opening the Black Box: Automated Decisions and the GDPR" SSRN Electronic Journal, 2017

⁹⁶ Gerke, Sara, et al. "Ethical and Legal Challenges of Artificial Intelligence-Driven Healthcare." Artificial Intelligence in Healthcare, 2020, pp. 295–336.

È probabile che le aziende che sono titolari del trattamento ai sensi del GDPR debbano effettuare una valutazione d'impatto sulla protezione dei dati per le nuove tecnologie basate sull'IA che saranno implementate nello spazio clinico.

L'articolo 35, paragrafo 3, del GDPR stabilisce esplicitamente quando è richiesta una valutazione d'impatto sulla protezione dei dati, ad esempio nei casi di "una valutazione sistematica e approfondita degli aspetti personali relativi alle persone fisiche basata sul trattamento automatizzato, compresa la profilazione, e su cui si basano decisioni che producono effetti giuridici riguardanti la persona fisica o incidono in modo analogo significativamente sulla persona fisica" o "trattamento su larga scala di categorie speciali di dati" (ad esempio, dati genetici e dati relativi alla salute).

L'articolo 35, paragrafo 7, del GDPR contiene un elenco di ciò che la valutazione deve almeno includere, come una descrizione dei trattamenti previsti, una valutazione dei rischi per le libertà e i diritti degli interessati e le misure previste per affrontare i rischi.

È difficile immaginare usi medici dell'IA che non implicino dati sanitari. Pertanto, l'IA medica dovrebbe essere utilizzata a causa del consenso dell'interessato o secondo una normativa specifica. Anzi, si potrebbe sottolineare la realistica autonomia dell'IA nell'ordinamento giuridico (italiano e) europeo.

Considerando, ad esempio, la legge italiana, l'articolo 2-sexies del Codice della Privacy⁹⁷ individua i motivi di interesse pubblico richiesti all'articolo 9, comma 2, lettera g) GDPR, per il trattamento di dati speciali con compiti di assistenza sanitaria nazionale, mentre l'art. 2-septies dello stesso Codice Privacy⁹⁸ descrive le misure idonee a tutelare i diritti, le libertà e gli interessi legittimi dell'interessato.

Analogamente, anche altri Stati membri dell'UE hanno previsto norme specifiche per il trattamento di dati speciali.

Il consenso richiesto dall'articolo 22, paragrafo 4, del GDPR, appare come l'approvazione e la legittimazione della decisione medica autonoma e autonoma. La necessità di un equilibrio tra i benefici dell'IA (ad esempio efficienza e precisione) e la sua "opacità" porterà l'interessato, cioè il paziente, a fare una scelta: da un lato, potrebbe beneficiare di un servizio intelligente, non

⁹⁷ Articolo 2-sexies Nuovo Codice Privacy – D.lgs 196/2003 agg. D.lgs 101/2018 CyberLaws. www.cyberlaws.it/2018/art-2-sexies-d-lgs-196-2003.

⁹⁸ Art. 2 septies codice della privacy – “Misure di garanzia per il trattamento dei dati genetici, biometrici e relativi alla salute”. Brocardi.it, si veda www.brocardi.it/codice-della-privacy/parte-i/titolo-i/capoi/art2septies.html.

comprendendo forse ciò che l'IA suggerisce; dall'altro, potrebbe rifiutare l'uso delle nuove tecnologie, magari perdendo l'innovatività dell'assistenza medica IA.

Pertanto, il paziente-interessato dovrà acconsentire al trattamento dei dati personali e al servizio medico (il consenso informato).

La giurisprudenza ha formulato il principio di diritto secondo cui: "in tema di trattamento di dati personali, il consenso è validamente prestato solo se espresso liberamente e specificamente in riferimento a un trattamento chiaramente individuato". In sostanza non può non può esserci consenso al trattamento automatizzato dei propri dati personali senza trasparenza dell'algoritmo.

La Corte di legittimità, con ordinanza n. 14381/2021, ha stabilito che 'algoritmo di funzionamento deve essere conosciuto ed oggetto di un consenso specifico da parte del cliente.

La Cassazione ha accolto il ricorso del Garante privacy nei confronti di un'associazione che si proponeva di contrastare "fenomeni basati sulla creazione di profili artefatti o inventieri e di calcolare in maniera imparziale il cd. rating reputazione, in modo da consentire a terzi una verifica di reale credibilità".⁹⁹

Nel prossimo futuro, forse, la tecnologia dell'informazione potrà essere implementata per personalizzare le divulgazioni al fine di tenere conto dell'eterogeneità degli attori e migliorare continuamente la pertinenza delle informazioni fornite.

L'uso dell'IA implica che l'utente riceva informazioni chiare e comprensibili sul funzionamento del sistema basato sull'IA in merito al trattamento dei dati personali e sulle modalità specifiche di interazione del sistema stesso con gli esseri umani. In effetti, la formazione del consenso nell'area specifica di indagine crea molte difficoltà.

Cercando di individuare una soluzione, in funzione di questa specifica finalità dell'IA, si propone di fare riferimento alla normativa sulla sperimentazione dei dispositivi medici e sulla ricerca biomedica, al fine di 'riportare' il rischio dato dall'IA all'interno della sfera del rischio sanitario connesso alla ricerca e alla sperimentazione medico-scientifica. Questa interpretazione potrebbe essere corretta per almeno due motivi.

Prima di tutto, c'è un argomento pratico: l'IA in sanità può essere qualificata come dispositivo medico quando è caratterizzata da uno degli usi e delle finalità dell'articolo 2 MDR (Regolamento europeo sui dispositivi medici).

⁹⁹ “*Rating reputazionale su internet, algoritmo in chiaro e consenso specifico*” 2021, si veda https://ntplusdiritto.ilsole24ore.com/art/rating-reputazionale-internet-algoritmo-chiaro-e-consenso-specifico-AE7DdqL?refresh_ce=1

In questo modo, un'interpretazione analogica sembra adattarsi al caso specifico dello studio.

Il secondo motivo, invece, è legato all'utilizzo di dati personali (particolari). L'autoapprendimento dell'IA, infatti, porta il sistema a trarre ispirazione da ogni caso già analizzato per alimentare e calibrare il proprio lavoro. Questa capacità di autoapprendimento autonomo, quindi, aiuta l'IA ad affrontare casi pratici non noti o sconosciuti nello stesso modo in cui un ricercatore utilizzerebbe le proprie conoscenze passate per studiare e risolvere sfide mediche sconosciute. In questa prospettiva, l'applicazione della normativa sulla ricerca e sui dispositivi medici può diventare uno strumento utile in grado di informare il paziente sui potenziali rischi dell'utilizzo dell'IA e lasciare alla sua volontà la scelta tra un trattamento sanitario autonomo e uno ordinario. Infatti, deve esistere un rapporto di proporzionalità tra lo scopo da raggiungere e i rischi potenzialmente imputabili al trattamento in questione (ai sensi dell'art. 15 e ss. della Convenzione di Oviedo).

Le peculiarità del caso specifico, tuttavia, possono indurre l'interprete a considerare altre regole dovute alla particolare modalità operativa del software o al fatto che il suo uso specifico è diventato di routine (tanto da essere incluso nelle linee guida e nei protocolli medici) e quindi ha raggiunto notevoli standard di sicurezza.

In generale, l'articolo 35, paragrafo 1, del GDPR richiede tale valutazione, prima del trattamento, per le "nuove tecnologie" in cui il trattamento "può comportare un rischio elevato per i diritti e le libertà delle persone fisiche".

Da questa analisi del GDPR sembra chiaro che il sistema giuridico non è, attualmente, preparato alle sfide poste dalle nuove tecnologie, soprattutto nella configurazione dei profili di responsabilità. Tuttavia, è possibile individuare criteri ermeneutici utili non solo agli studiosi, ma anche agli operatori pratici.¹⁰⁰

È chiaro che la legislazione sulla protezione dei dati non può regolare ogni aspetto dell'uso dell'IA, ma allo stesso tempo il GDPR sembra avere una carica personalistica, che consente di applicarlo oltre i normali confini della privacy.

Il GDPR, inoltre, regola ampiamente la dignità della persona umana, e la privacy è certamente uno dei suoi aspetti.

Da questo punto di vista, il primo strumento ermeneutico può essere trovato nel modello di accountability.

¹⁰⁰ Ciancimino M., "AI-Based Decision-Making Process in Healthcare", (2022), 11, *Journal of European Consumer and Market Law*, Issue 5, pp. 173-180, <https://kluwerlawonline.com/JournalArticle/Journal+of+European+Consumer+and+Market+Law/11.5/EuCML2022029>

Il Gruppo europeo di etica descrive la “responsabilità morale” come un aspetto integrante “della concezione della persona su cui si fondano tutte le nostre istituzioni morali, sociali e giuridiche”. Il concetto è inoltre inteso come la capacità di “essere responsabile sia che qualcosa riesca o fallisca”¹⁰¹. Il termine combina diversi aspetti dell'agire umano e deve essere considerato come “obbligo di rendere conto”. Approssimativamente, ciò comprende almeno due diverse categorie, quella della giustificazione nel senso di spiegabilità delle proprie azioni¹⁰¹ e quella del risarcimento, compreso il risarcimento e l'attribuzione delle colpe.

La responsabilità dei produttori, degli sviluppatori e degli utenti può portare, in primo luogo, all'attuazione di processi di follow-up dei sistemi di IA, come i dispositivi medici.¹⁰²

In secondo luogo, può portare a valutazioni sulla protezione dei dati anche durante l'uso professionale dell'algorithm autonomo.

Questo approccio potrebbe essere utile perché ogni sistema, grazie alle sue capacità di autoapprendimento, ha un certo grado di unicità e originalità. Inoltre, queste qualità richiedono un'analisi approfondita, sia giuridica che etica, e, dal punto di vista della protezione dei dati, un approccio caso per caso (o, almeno, settore per settore).

Quest'ultimo dovrebbe prendere in considerazione le specificità del singolo sistema di IA, come i tipi di dati trattati, le modalità di acquisizione dei dati, la capacità di autoapprendimento del sistema, ecc.

Inoltre, potrebbe essere consigliabile una valutazione delle norme giuridiche pertinenti in una prospettiva di diritto personalizzato.

L'elaborazione delle informazioni attraverso l'IA potrebbe consentire di "personalizzare" la legge e raggiungere un livello di "granularità" che non è stato ancora raggiunto. Forse c'è la possibilità di ripensare il rapporto tra il sistema giuridico e l'IA stessa, considerando questa tecnologia non solo come uno strumento, ma come un oggetto singolare delle regole legali che necessita di una regolamentazione personalizzata.

Le ragioni sopra menzionate evidenziano l'importanza di un'analisi approfondita dell'IA nell'assistenza sanitaria, alla luce della necessità intrinseca di attuare e aggiornare i valori e principi

¹⁰¹ V Dignum, ‘*The ART of AI — Accountability, Responsibility, Transparency, Medium*’, 2019, <https://medium.com/@virginiadignum/the-art-of-ai-accountability-responsibility-transparency-48666ec92ea5>.

¹⁰² Gerke, Sara, et al. “*Ethical and Legal Challenges of Artificial Intelligence-Driven Healthcare.*” *Artificial Intelligence in Healthcare*, 2020, pp. 295–336., <https://doi.org/10.1016/b978-0-12-818438-7.00012-5>.

dell'ordinamento giuridico dell'UE. In tal modo, è possibile individuare un obiettivo interessante nella cooperazione tra AI e operatori sanitari.

La necessità di evitare che l'essere umano diventi protesi di una macchina impone all'interprete di attuare la riflessione – soprattutto in questo momento storico in cui le innovazioni tecnologiche stanno fortemente sollecitando sia le categorie giuridiche che quelle antropologiche – e, per quanto possibile, di lavorare in sinergia con la più ampia comunità scientifica al fine di sostenere il progresso nella ricerca del giusto equilibrio tra esigenze di tutela personalistica e richieste di innovazione. e miglioramento della qualità della vita. In conclusione, quindi, la cooperazione è la chiave.

Sintetizzando, in ambito sanitario, si dovrebbe fare attenzione al fatto che gli oggetti del trattamento sono (formalmente o sostanzialmente) categorie particolari di dati personali sub specie dati sanitari che richiedono un livello di protezione più elevato. Al fine di valutare il ruolo dell'autonomia nel processo decisionale sanitario, ci si dovrà e concentrarsi sul tema (correlato) del "riutilizzo" dei dati personali.

L'IA, come conseguenza del suo autoapprendimento, trae continuamente ispirazione dai dati precedentemente analizzati per la futura elaborazione dei dati. Dall'analisi finora effettuata sembra condivisibile ritenere che il riutilizzo dei dati (vale a dire, il trattamento dei dati per ulteriori scopi) sia lecito.

Si aprono così due scenari principali dell'elaborazione dei dati ICT in medicina.

Il primo scenario è quando un software viene utilizzato – sotto il quasi pieno controllo di un essere umano – direttamente per uso terapeutico e a) è richiesto un consenso ad hoc per il trattamento dei dati relativi alla salute, o b), esiste una base giuridica per il trattamento dei dati, perché questo tipo di utilizzo è tra quelli per i quali non è richiesto il consenso ai sensi dell'articolo 9, paragrafo 2, del GDPR. Più specificamente, il riutilizzo dei dati è consentito ai sensi dell'articolo 5, paragrafo 1, lettera b), del GDPR. Tale provvedimento, infatti, autorizza espressamente "ulteriori trattamenti" perché compatibile con le finalità per le quali i dati sono stati principalmente raccolti, e sembra che l'uso terapeutico sia compatibile con l'utilizzo precedente, ovvero la finalità di trattamento medico per la quale i dati stessi sono stati acquisiti.

La stessa soluzione sembra essere ammissibile per il secondo scenario più sofisticato in cui il trattamento dei dati è privo di un pieno controllo umano, vale a dire nell'ipotesi dell'articolo 22, paragrafo 4, del GDPR: ciò vale quando la base giuridica è il consenso ai sensi dell'articolo 9, paragrafo 2, lettera a), del GDPR e nel caso del trattamento dei dati di cui all'articolo 9, paragrafo 2, lettera g), vale a dire "motivi di interesse generale fondati sul diritto dell'Unione o degli Stati

membri", adottando ovviamente misure specifiche per proteggere i diritti e le libertà fondamentali delle persone.

E' condivisibile l'orientamento secondo cui la liceità del riutilizzo dei dati sanitari dipende dalla compatibilità tra le finalità per le quali i dati sono ulteriormente trattati.

In realtà, il riutilizzo attraverso l'AI può avere uno scopo (compatibile con quello principale che ne ha permesso l'uso primario) di terapia, perché il software intelligente elabora gli stessi dati precedentemente raccolti al fine di migliorare i successivi usi terapeutici o terapeutici. Pertanto, la "volontà" dell'IA di migliorare le sue capacità terapeutiche dovrebbe dare uno scudo di legalità all'ulteriore elaborazione dei dati.

Il riutilizzo dei dati effettuato da un sistema sanitario autonomo potrebbe anche essere qualificato come lecito ai fini della ricerca scientifica. In questo modo, ad esempio, l'espansione del dataset consente all'AI di studiare nuovi modelli e trovare nuovi collegamenti semantici al fine di valutare e formulare ipotesi diagnostiche o terapeutiche innovative.

6. Il Regolamento della Commissione europea sull' I.A.: l' Artificial Intelligence Act e il modello basato sul rischio

E' notoriamente difficile tentare di normare la tecnologia.

Il principio di neutralità tecnologica ormai affermato a livello internazionale è il risultato di un dibattito complesso. Secondo questo principio affermato durante lo sviluppo di UNCITRAL (Commissione delle Nazioni Unite per il diritto commerciale internazionale) ¹⁰³, il diritto deve essere tecnologicamente neutrale, evitando l'affermazione del principio di determinare soluzioni tecnologiche che consentano l'esecuzione delle procedure giudiziarie.

In altre parole, le norme giuridiche non dovrebbero fare riferimento a un livello di sicurezza predeterminato o a una specifica tecnologia, ma dovrebbero limitarsi a indicare ciò che si vuole ottenere senza specificare i mezzi tecnici per raggiungerlo" ¹⁰⁴.

¹⁰³ Ratti, M. "La Convenzione sull'uso delle Comunicazioni Elettroniche: Le principali disposizioni.", in *Diritto dell'informatica*, 2017, p. 71, si veda <https://hdl.handle.net/11585/580169>.

¹⁰⁴ Il principio di neutralità tecnologica è stato adottato dal legislatore europeo, ad esempio, nel Regolamento (UE) n. 910/2014, il c.d. Reg. e-IDAS, allorché ha introdotto delle definizioni di firma elettronica e firma

Quando si regola l'IA, devono essere considerate altre due questioni chiave che possono limitare le azioni dei legislatori. Il primo è incorniciato dalla paura: la paura che applicazioni di intelligenza artificiale che potrebbero prendere decisioni in modo autonomo e, ispirate alla fantascienza, possano agire contro gli umani in modi inaspettati. Alla radice di questi timori c'è la difficoltà di prevedere l'esito del processo di machine learning.

Le proposte di regolamento europeo sull'intelligenza artificiale si inseriscono in un disegno strategico e si raggruppano essenzialmente in quattro ambiti: la protezione dei dati personali, ampiamente discussa; i servizi digitali e i mercati digitali, il Digital Services Act ¹⁰⁵ e il Digital Markets Act ¹⁰⁶; con la revisione del regolamento e-IDAS nel 2014 ¹⁰⁷.

L'obiettivo strategico è garantire la creazione di un mercato unico digitale europeo e, sviluppando un modello normativo, che l'UE "diventi un leader mondiale nello sviluppo di un'IA sicura, protetta ed etica" ¹⁰⁸. Gli obiettivi specifici sono i seguenti:

- assicurare che i sistemi di IA immessi sul mercato dell'Unione e utilizzati siano sicuri e rispettino la normativa vigente in materia di diritti fondamentali e i valori dell'Unione;
- assicurare la certezza del diritto per facilitare gli investimenti e l'innovazione nell'intelligenza artificiale;
- migliorare la governance e l'applicazione effettiva della normativa esistente in materia di diritti fondamentali e requisiti di sicurezza applicabili ai sistemi di IA;
- facilitare lo sviluppo di un mercato unico per applicazioni di IA lecite, sicure e affidabili nonché prevenire la frammentazione del mercato”.

La proposta recente del 2021 per regolamentare l'intelligenza artificiale è una proposta molto complessa con 85 articoli, dal titolo “Laying Down Harmonised Rules On Artificial Intelligence (**Artificial Intelligence Act**) And Amending Certain Union Legislative Acts”.¹⁰⁹

elettronica avanzata non riferite a specifiche tecnologie. Non può dirsi, invece, “tecnologicamente neutra” la nozione di firma elettronica qualificata, collegata a livelli di sicurezza predeterminati.

¹⁰⁶ Proposta di Regolamento del Parlamento europeo e del Consiglio relativo a mercati equi e contendibili nel settore digitale (legge sui mercati digitali), 2020, consultabile al <https://eur-lex.europa.eu/legal-content/it/TXT/?uri=COM:2020:842:FIN>

¹⁰⁷ Proposta di Regolamento del Parlamento europeo e del Consiglio che modifica il regolamento (UE) n. 910/2014 per quanto riguarda l'istituzione di un quadro per un'identità digitale europea, 2021, consultabile al <https://eur-lex.europa.eu/legal-content/IT/TXT/?uri=CELEX:52021PC0281>.

¹⁰⁸ così la Relazione della Commissione europea che accompagna la proposta di Regolamento.

Secondo la Commissione, è stata scelta una metodologia che prevede un quadro normativo rivolto solo ai sistemi di IA ad alto rischio, con la possibilità per tutti i fornitori di sistemi di IA a basso rischio di seguire un codice di condotta. "I requisiti riguardano dati, documentazione e tracciabilità, divulgazione e trasparenza delle informazioni, controllo umano, affidabilità e accuratezza, e sono obbligatori per i sistemi di intelligenza artificiale ad alto rischio. Le aziende che li implementano lo faranno volontariamente"¹¹⁰.

Dunque, la proposta disciplina l'immissione nel mercato, la messa a disposizione, la messa in servizio e l'uso dei sistemi di intelligenza artificiale, come recita l'art. 1, in generale, e non in un settore specifico.

Secondo questo approccio, la definizione di intelligenza artificiale è molto generica.

Il modello adottato dalla Commissione europea è un modello basato sul rischio che classifica l'uso dell'IA rispettivamente in rischi inaccettabili, alti, bassi o minimi, come descritto nello stesso rapporto sulla proposta di regolamento.

I legislatori stabiliscono che esistono applicazioni di intelligenza artificiale che dovrebbero essere vietate e vengono elencate all'art. 5:

- 1) sistemi di IA che adoperino tecniche subliminali oltre la coscienza di una persona per distorcerne il comportamento materiale in una maniera che causi o crei il rischio di arrecare un danno fisico o psicologico a sé o a un terzo;
- 2) sistemi IA che sfruttano le vulnerabilità di uno specifico gruppo di persone collegate all'età, alla disabilità fisica o mentale, per influenzare il comportamento di una persona appartenente a tale gruppo in una maniera che causi o crei il rischio di arrecare un danno fisico o psicologico a sé o a un terzo;
- 3) sistemi IA utilizzati dalle pubbliche autorità o a beneficio di quest'ultime, per la valutazione o classificazione della affidabilità di una persona fisica in un determinato periodo di tempo, sulla base del suo comportamento sociale ovvero delle sue caratteristiche personali o di personalità note o previste.¹¹¹

In particolare, per quanto riguarda queste pratiche di IA, i legislatori europei hanno incluso la performance oggettiva come un obiettivo che caratterizza le applicazioni vietate.

È altresì vietato vendere o mantenere sistemi di conteggio sociale lesivi della dignità umana.

¹¹¹ Yüksel, A. E. B. "A General Look at the Proposal for a Regulation on Artificial Intelligence of the European Commission" *Türkiye Adalet Akademisi Dergisi*, 2022, pp. 19–46

Infine, è vietato anche l'uso di sistemi biometrici remoti “in tempo reale” in luoghi pubblici, soprattutto per finalità di contrasto che violino i diritti e le libertà degli interessati.

L'importanza, ovviamente, risiede nelle eccezioni previste dal legislatore. Ad esempio, se l'utilizzo di un sistema biometrico remoto è strettamente necessario per la “ricerca mirata di specifiche potenziali vittime di reato, compresi i minori scomparsi” o per la “prevenzione del crimine”, ecc. integrità di un individuo, o minacci ad atto terroristico” allora non può essere considerato nel novero delle applicazioni vietate.

Si tratta di una discussione aperta, in quanto deroghe troppo ampie possono invalidare quanto vietato da altre disposizioni.

La proposta di Regolamento europeo, dall'art. 6 in poi, classifica i sistemi di intelligenza artificiale ad alto rischio e prevede che se i sistemi di intelligenza artificiale sono sistemi ad alto rischio, allora il soggetto che li produce e che li immette nel mercato deve adempiere a una serie di prescrizioni e procedere ad una valutazione della conformità ex ante.

Sono sistemi classificati ad alto rischio: i sistemi destinati ad essere utilizzati come componenti di sicurezza di prodotti soggetti a valutazione di conformità ex ante da parte di terzi; altri sistemi di IA indipendenti, che presentano implicazioni in relazione ai diritti fondamentali, elencati all'allegato III del Regolamento.

Gli obblighi e i requisiti cui deve conformarsi un sistema di intelligenza artificiale ad alto rischio sono dettati invece dall'art. 8 e seguenti.

Ai sensi dell'art. 8, per garantire il rispetto di tali obblighi e devono tener conto di due elementi: il sistema di IA ad alto rischio e il sistema di gestione dei rischi, soggetti alla disciplina specificatamente dettata dall'art. 9.

Innanzitutto, è necessario stabilire, implementare, documentare e mantenere un sistema di gestione, vale a dire un processo iterativo continuo, eseguito durante tutto il ciclo di vita del sistema di IA ad alto rischio, compresa l'identificazione e analisi dei rischi noti e prevedibili per il sistema, stima e valutazione dei rischi che possono verificarsi quando il sistema di IA è ad alto rischio utilizzato per il suo scopo e in condizioni avverse ragionevolmente prevedibili, valutazione di ogni altro rischio derivante dai dati raccolti dal sistema di sorveglianza post-vendita, nonché che l'adozione di adeguate misure di gestione del rischio.¹¹²

¹¹² Ebers, M., Hoch, V. R., Rosenkranz, F., Ruschemeier, H., & Steinrötter, B. “*The European Commission’s proposal for an artificial intelligence act—a critical assessment by members of the Robotics and AI Law Society*” 2021, 589–603 si veda <https://doi.org/10.3390/j4040043>.

Successivamente, alcuni criteri di qualità sono set di dati identificati utilizzati per la convalida e per scopi di test nel contesto dell'IA ad alto rischio.

In particolare, questi dataset devono essere soggetti a pratiche di governance e gestione specifiche e pertinenti, rappresentative, prive di errori, complete e statisticamente appropriate.

Inoltre, è necessario che prima di immettere sul mercato o di mettere in servizio l'IA ad alto rischio, la relativa documentazione tecnica sia conforme ai requisiti stabiliti dal regolamento e fornisca alle autorità e agli organismi competenti tutte le informazioni per valutare tale conformità.¹¹³

Inoltre, i sistemi di IA ad alto rischio devono essere progettati e sviluppati in modo tale da garantire, attraverso eventi automatici e durante tutto il ciclo di vita, la tracciabilità del loro funzionamento, che dovrebbe essere sufficientemente trasparente per consentire agli utenti di interpretare l'output da utilizzare in modo appropriato.

In quest'ottica, i sistemi di IA ad alto rischio devono essere accompagnati da istruzioni per l'uso, in formato digitale o contenenti informazioni concise, complete, corrette, chiare e pertinenti e comprensibili per gli utenti.

I sistemi di intelligenza artificiale ad alto rischio devono quindi essere progettati e sviluppati in modo tale da conseguire, in relazione al loro scopo, un adeguato grado di precisione, robustezza e sicurezza informatica, che lungo tutto il ciclo di vita del sistema.

Altri obblighi e requisiti per i sistemi di IA a rischio derivano da disposizioni che stabiliscono obblighi specifici per i soggetti coinvolti nella catena della realizzazione e di tali sistemi.

Si tratta del fornitore del sistema di IA, definito come persona fisica o giuridica, autorità pubblica, agenzia o qualsiasi organismo che sviluppa o ha sviluppato un sistema di intelligence al fine di collocarlo sul mercato con il proprio nome o con il proprio marchio, a titolo oneroso o dell'utente, vale a dire qualsiasi persona fisica o persona giuridica, autorità di agenzia o altro organismo che utilizza, per scopi non personali, un sistema di intelligenza artificiale posto sotto autorità.

L'importatore è la persona fisica o giuridica stabilita in chi immette sul mercato o mette in servizio un sistema di intelligenza artificiale recante il nome o il marchio di una persona fisica o giuridica stabilita al di fuori dell'Unione;

¹¹³ Ebers, M., Hoch, V. R., Rosenkranz, F., Ruschemeier, H., & Steinrötter, B. “*The European Commission’s proposal for an artificial intelligence act—a critical assessment by members of the Robotics and AI Law Society*” 2021, 589–603 si veda <https://doi.org/10.3390/j4040043>.

Il *distributor* è qualsiasi persona fisica o giuridica nella catena di approvvigionamento, diversa dal fornitore o dall'importatore, che mette a disposizione un sistema di intelligenza artificiale sul mercato dell'Unione senza modificarne le proprietà .

Nel quadro delineato nella proposta, i fornitori di sistemi di IA ad alto rischio sono innanzitutto tenuti a garantire che i sistemi siano conformi ai requisiti già elencati, nonché ad adottare immediatamente le misure correttive necessarie in caso di discrepanze.

Anche la conformità del sistema di IA ad alto rischio alle normative deve essere certificata da una procedura di valutazione, il fornitore ne garantisce il raggiungimento prima che venga immesso sul mercato o messo in servizio.

Se la procedura di valutazione della conformità ha esito positivo, deve redigere una dichiarazione di conformità e apporre la CE.

In tutti i casi, il fornitore dovrà dimostrare la conformità del sistema di IA ad alto rischio con il regolamento sui requisiti quando richiesto da un'autorità nazionale competente.

Inoltre istituisce e documenta un sistema di monitoraggio che, proporzionato alla natura delle tecnologie e poste dal sistema di IA, gli consente di valutarne la continuità.

Anche in questo caso, il fornitore di servizi è tenuto a implementare un sistema di gestione della qualità attraverso politiche, procedure e istruzioni scritte; a garantire la conformità del sistema di gestione della qualità IA alle regole; a redigere la documentazione tecnica del sistema; a tenere sotto il proprio controllo i log generati automaticamente prima che il sistema di IA ad alto rischio sia immesso sul mercato o messo in servizio.

Per quanto riguarda gli importatori, si prevede che prima di immettere sul mercato un sistema di IA ad alto rischio, essi devono verificare che il fornitore abbia svolto l'opportuna procedura di valutazione della conformità e redatto la relativa documentazione tecnica.

Gli importatori assicurano inoltre che il sistema sia provvisto del necessario certificato di conformità e sia accompagnato dalla necessaria documentazione di istruzioni operative.

Inoltre, qualora l'importatore ritenga o abbia motivo di ritenere che un sistema di IA ad alto rischio non sia sicuro, egli è tenuto a non immetterlo sul mercato finché non abbia non è stato reso conforme e, se il sistema presenta un rischio per la salute, per la sicurezza o per la protezione delle persone fisiche, deve informare il fornitore e le autorità di vigilanza del mercato.

Analoghi obblighi incombono, a cascata, sul distributore che, per rendere disponibile sul mercato un sistema ad alto rischio, non solo controlla che il sistema abbia la marcatura CE, il documentazione tecnica e istruzioni ma, in generale, che il fornitore e l'importatore rispettino gli obblighi del regolamento.

Ciò comporta gli obblighi degli utenti, tenuti a utilizzare l'IA ad alto rischio in conformità con le istruzioni per l'uso, il funzionamento del sistema e a informare il fornitore o il distributore, oltre a sospendere l'uso del sistema, quando hanno motivo di ritenere che questo presenti un pericolo per la salute, la sicurezza o per i diritti fondamentali delle persone o se riscontrino gravi malfunzionamenti.

Giova rilevare che in virtù dell' art. 28¹¹⁴, qualsiasi distributore, importatore, utilizzatore o altra terza parte è assimilato a un fornitore ai fini della normativa, ed è soggetto agli obblighi ad essa relativi, se immette sul mercato un sistema di IA ad alto rischio con il proprio nome o marchio; se modifica ad alto la destinazione di un sistema di IA che è già stato immesso sul mercato o in servizio e, infine, se apporta una modifica sostanziale al sistema di IA ad alto rischio.

Nonostante l'innovativa struttura basata sul rischio, l'AIA sembra presentare notevoli lacune.

L'AIA - e questo è senza dubbio uno dei punti più cruciali - non prevede alcun diritto individuale. Sebbene il regolamento sia inteso a tutelare i diritti fondamentali, manca di rimedi attraverso i quali i singoli possano chiedere riparazione per una violazione del regolamento.

In particolare, il progetto non prevede alcun meccanismo per facilitare il ricorso degli individui contro il processo decisionale guidato dall'IA, ma appare necessario che gli diritti individuali siano ben coordinati con i diritti esistenti.¹¹⁵

E' evidente che l'AIA miri ad attuare gli obiettivi di politica legale in conformità con i valori europei. Tuttavia, un approccio europeo nei confronti dell'IA dovrebbe rispettare non solo i diritti umani e i valori etici, ma anche altri obiettivi come il cambiamento climatico e la sostenibilità. A questo proposito, l'AIA non menziona direttamente "Green AI" o "Sustainable AI" come obiettivo concreto di una comprensione europea dello sviluppo di un'IA a prova di futuro e sostenibile, che sono anche temi importanti del Green Deal europeo¹¹⁶.

La proposta riconosce che l'IA può sostenere risultati socialmente e ambientalmente vantaggiosi e che tale azione è necessaria nel settore ad alto impatto del cambiamento climatico. L'AIA afferma che il diritto fondamentale a un elevato livello di protezione ambientale dovrebbe essere preso in

¹¹⁴ Yüksel, A. E. B. "A General Look at the Proposal for a Regulation on Artificial Intelligence of the European Commission" *Türkiye Adalet Akademisi Dergisi*, 2022 pp. 19–46.

¹¹⁶ Gailhofer, P.; Herold, A.; Schemmel, Giappone; Scherf, CS-S.; Urruzia, C.; Kohler, A.; Braungardt, S. *The Role of Artificial Intelligence in the European Green Deal*, studio richiesto dalla commissione AIDA del Parlamento europeo, 2021 consultabile al [https://www.europarl.europa.eu/RegData/etudes/STUD/2021/662906/IPOL_STU\(2021\)662906_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2021/662906/IPOL_STU(2021)662906_EN.pdf)

considerazione nel valutare il potenziale danno dei sistemi di IA. Ciononostante, ci sono solo disposizioni sparse riguardanti la protezione dell'ambiente, ad esempio, ai sensi dell'art. 62, i fornitori di sistemi di IA ad alto rischio hanno l'obbligo di segnalare il danno grave all'ambiente.

Tutto sommato, nonostante alcuni difetti e chiarimenti necessari, la proposta della Commissione europea per una legge sull'intelligenza artificiale ha molti elementi innovativi ed è generalmente ben progettata.

L'approccio normativo basato sul rischio non è solo appropriato, ma anche il più pratico. Tuttavia, un aspetto principale che necessita di miglioramenti è la definizione del termine "AI".

L'AIA utilizza una definizione piuttosto ampia e crea quindi il rischio di un'eccessiva regolamentazione dei sistemi, che in genere non rappresentano alcun pericolo derivante dalle tecniche di riciclaggio e intelligenza artificiale.

La seconda critica principale riguarda la proposta struttura di (auto)applicazione, che si basa principalmente su controlli interni condotti dal fornitore di IA.

Nella maggior parte dei casi, l'AIA non richiede alcuna supervisione esterna. Inoltre, le disposizioni relative alle norme e ai criteri di conformità appaiono spesso frammentarie e imprecise.

Indipendentemente da queste critiche è possibile credere che l'AIA sia una misura appropriata per creare un ambiente normativo sicuro e affidabile per l'IA in Europa e, che quindi, sia un passo nella giusta direzione.

Le definizioni degli obblighi degli operatori dell' IA delineate dall' AIA saranno utili per ricostruire un excursus sotto il profilo della responsabilità per la riparazione del danno cagionato dalle macchine.

7. OMS: Raccomandazioni e linee guida in merito allo sviluppo, all'applicazione e all'uso delle tecnologie di intelligenza artificiale (Artificial Intelligence, AI) nel settore della medicina.

L'Organizzazione Mondiale della Sanità ha elaborato un insieme di indicazioni, raccomandazioni e linee guida in merito allo sviluppo, all'applicazione e all'uso delle tecnologie di intelligenza artificiale (Artificial Intelligence, AI) nel settore della medicina.

Il rapporto Etica e governance dell'intelligenza artificiale per la salute, prodotto da OMS nel giugno 2021, analizza le opportunità e le principali sfide poste dall'avvento dell'intelligenza artificiale in

campo medico, evidenziandone i vantaggi e i rischi, raccomanda lo sviluppo di politiche, principi e pratiche per l'uso etico di l'intelligenza artificiale in medicina e individua gli strumenti da utilizzare per evitare l'uso improprio dell'IA che potrebbe ledere la tutela dei diritti fondamentali.

Il documento è composto da nove sezioni ed un annesso.¹¹⁷

La prima sezione illustra le ragioni dell'impegno di World Health in questo contesto e indica i principali destinatari dei risultati e delle indicazioni contenute nel documento.

L'OMS osserva che, ad oggi, non esiste una guida internazionale completa sull'uso dell'IA per la salute in conformità con le norme etiche e gli standard sui diritti umani. L'organizzazione riconosce che la guida etica basata sulle prospettive condivise delle diverse entità che sviluppano, utilizzano o supervisionano tali tecnologie è fondamentale per creare fiducia in queste tecnologie e per proteggersi dagli effetti negativi. Sebbene i lettori principali di questo documento di orientamento siano i ministeri della salute, potrebbe anche essere rilevante per altre agenzie governative, che utilizzano le tecnologie di intelligenza artificiale per la salute e le entità che progettano e finanziano le tecnologie di intelligenza artificiale per la salute.

La seconda e la terza sezione mirano a definire ulteriormente il contesto applicativo dell'intelligenza artificiale.

In particolare, la seconda sezione propone una definizione delle diverse tecniche di machine learning utilizzate da questi sistemi e una definizione di medical big data.

La terza sezione, invece, identifica le principali applicazioni dell'IA in medicina, comprese le applicazioni utilizzate nei paesi meno sviluppati, l'uso dell'intelligenza nella ricerca sanitaria, lo sviluppo di farmaci, la gestione e pianificazione dei sistemi sanitari e la sorveglianza della sanità pubblica. Per citare un esempio: gli strumenti di intelligenza artificiale sono stati utilizzati per guidare l'allocazione delle scarse risorse sanitarie durante la pandemia di COVID-19, come i vaccini, per i quali la fornitura iniziale era insufficiente.

Per un programma nazionale per vaccinare milioni di persone contro la SARS-Cov2, l'India, ad esempio, ha utilizzato il suo sistema nazionale di identificazione digitale, Aadhar, per evitare duplicazioni e per tracciare i beneficiari.¹¹⁸

La quarta sezione descrive poi le leggi, le discipline, le politiche e i principi che sono e saranno applicati in relazione all'uso di questa tecnologia nel contesto medico-sanitario, che include gli

¹¹⁸ Ethics and governance of artificial intelligence for health: WHO guidance - <https://www.who.int/publications/i/item/9789240029200>

strumenti messi in atto a tutela dei diritti fondamentali, le norme che disciplinano la protezione dei dati personali, le norme e le politiche relative all'utilizzo dei dati sanitari. Questi includono gli obblighi in materia di diritti umani che si applicano all'intelligenza artificiale, al ruolo delle leggi e ai quadri sulla protezione dei dati e altre leggi e politiche sui dati sanitari.

Inoltre, questa sezione presta attenzione anche alle differenze che fanno riferimento a principi etici che possono assumere rilevanza in relazione a l'uso dell'IA in campo medico, nonché ai diversi ruoli assunti da bioetica, diritto, ordine pubblico e dai vari quadri come fonti di standard etici in questo settore.

La sezione 5 descrive i sei principi etici che il gruppo di esperti ha identificato come guida per lo sviluppo e l'uso dell'IA per la salute. Loro includono:

1. Proteggere l'autonomia umana: l'uso dell'intelligenza artificiale può portare a situazioni in cui il potere decisionale potrebbe essere trasferito alle macchine. Nel contesto dell'assistenza sanitaria, il principio di autonomia significa che: le persone dovrebbero mantenere il controllo dei sistemi sanitari e delle decisioni mediche; i fornitori devono disporre delle informazioni necessarie per un uso sicuro ed efficace dei sistemi di intelligenza artificiale e i pazienti devono comprendere il ruolo che tali sistemi svolgono nella loro assistenza; la privacy e la riservatezza dovrebbero essere protette; e i pazienti devono dare un valido consenso informato attraverso quadri giuridici appropriati per la protezione dei dati.
2. Promuovere il benessere e la sicurezza umana e l'interesse pubblico: le tecnologie di intelligenza artificiale non dovrebbero danneggiare le persone. I progettisti di tecnologie di intelligenza artificiale dovrebbero soddisfare i requisiti normativi per la sicurezza, l'accuratezza e l'efficacia per casi d'uso o indicazioni ben definiti. Devono essere disponibili misure di controllo della qualità nella pratica e miglioramento della qualità nell'uso dell'IA.
3. Garantire trasparenza, spiegabilità e intelligibilità: la trasparenza richiede che vengano pubblicate o documentate informazioni sufficienti prima della progettazione o dell'implementazione di una tecnologia AI. Tali informazioni devono essere facilmente accessibili e facilitare consultazioni e dibattiti pubblici significativi su come è progettata la tecnologia e su come dovrebbe o non dovrebbe essere utilizzata.
4. Promuovere responsabilità: sebbene le tecnologie di intelligenza artificiale svolgano compiti specifici, è responsabilità delle parti interessate garantire che vengano utilizzate in condizioni adeguate e da persone adeguatamente formate. Dovrebbero essere disponibili meccanismi efficaci per l'interrogatorio e per la riparazione di individui e gruppi che sono influenzati negativamente da decisioni basate su algoritmi.

5. Garantire l'inclusione e l'equità: l'inclusione richiede che l'IA per la salute sia progettata per incoraggiare l'uso e l'accesso più equo possibile, indipendentemente da età, sesso, sesso, reddito, razza, etnia, orientamento sessuale, capacità o altre caratteristiche protette dai codici dei diritti umani . Gli strumenti e i sistemi di intelligenza artificiale dovrebbero essere monitorati e valutati per identificare effetti sproporzionati su gruppi specifici di persone. Nessuna tecnologia dovrebbe sostenere o peggiorare le forme esistenti di pregiudizio e discriminazione.
6. Promuovere un'IA reattiva e sostenibile: progettisti, sviluppatori e utenti dovrebbero valutare in modo continuo e trasparente le applicazioni di intelligenza artificiale durante l'uso effettivo per determinare se l'IA risponde in modo adeguato e appropriato alle aspettative e ai requisiti. L'uso dell'intelligenza artificiale dovrebbe essere coerente con gli sforzi globali per ridurre l'impatto degli esseri umani sull'ambiente, sugli ecosistemi e sul clima della Terra. I governi e le aziende dovrebbero considerare e affrontare questioni come la potenziale perdita di posti di lavoro dovuta all'uso di sistemi automatizzati.

La sezione 6 presenta le sfide etiche identificate e discusse dal gruppo di esperti a cui possono essere applicati questi principi etici guida: se l'IA debba essere utilizzata; IA e divario digitale; raccolta e utilizzo dei dati; responsabilità e responsabilità del processo decisionale con l'IA; processo decisionale autonomo; pregiudizio e discriminazione associati all'IA; rischi dell'IA per la sicurezza e la sicurezza informatica; impatti dell'IA sul lavoro e sull'occupazione nell'assistenza sanitaria; sfide nella commercializzazione dell'IA per l'assistenza sanitaria; e IA e cambiamento climatico. ¹¹⁹

Ad esempio, l'uso non regolamentato dell'intelligenza artificiale potrebbe subordinare i diritti e gli interessi dei pazienti e delle comunità ai potenti interessi commerciali delle società tecnologiche o agli interessi dei governi nella sorveglianza e nel controllo sociale.

Le sezioni conclusive del rapporto identificano misure legali, regolamentari e non legali per promuovere l'uso etico dell'IA per la salute, offrendo una serie di raccomandazioni.

La sezione 7 esamina come le varie parti interessate possono introdurre pratiche, programmi e misure etiche per anticipare o soddisfare le norme etiche e gli obblighi legali. Questa sezione include:

¹¹⁹ Rohit Malpani, and About World Health Organization (WHO) More knowledge from World Health Organization (WHO). “*Ethics and Governance of Artificial Intelligence for Health: Who Guidance.*” The Communication Initiative Network, 2021, si veda <https://www.comminit.com/global/content/ethics-and-governance-artificial-intelligence-health-who-guidance>.

la progettazione etica e trasparente delle tecnologie AI; i meccanismi per l'impegno e il ruolo del pubblico e dimostrazione di affidabilità con fornitori e pazienti; la valutazione di impatto; e un programma di ricerca per l'uso etico dell'IA per l'assistenza sanitaria.

Ecco alcuni esempi di raccomandazioni:

- Il pubblico dovrebbe essere coinvolto nello sviluppo dell'IA per la salute al fine di comprendere le forme di condivisione e utilizzo dei dati, commentare le forme di IA che sono socialmente e culturalmente accettabili ed esprimere pienamente le proprie preoccupazioni e aspettative. Inoltre, l'alfabetizzazione del pubblico in generale nella tecnologia dell'intelligenza artificiale dovrebbe essere migliorata per consentire loro di determinare quali tecnologie di intelligenza artificiale siano accettabili.
- Dovrebbero essere disponibili programmi di formazione e formazione continua per assistere gli operatori sanitari nella comprensione e nell'adattamento all'uso dell'IA, nell'apprendere i suoi benefici e rischi e comprendere le questioni etiche sollevate dal loro uso.

La sezione 8 discute in che modo i regimi di responsabilità possono evolversi con l'aumento dell'uso dell'IA per l'assistenza sanitaria, compreso il modo in cui la responsabilità potrebbe essere assegnata a un operatore sanitario, un fornitore di tecnologia e un sistema sanitario o un ospedale che seleziona una tecnologia di IA e come le regole di responsabilità potrebbe influenzare il modo in cui un professionista utilizza l'IA.

Questa sezione considera anche come risarcire le persone danneggiate dalle tecnologie AI e il ruolo delle agenzie di regolamentazione. E' raccomandabile che:

- Le agenzie internazionali e le società professionali garantissero che le loro linee guida cliniche tengano il passo con la rapida introduzione delle tecnologie di intelligenza artificiale attraverso l'apprendimento continuo;
- L'OMS dovrebbe supportare le agenzie di regolamentazione nazionali nella valutazione delle tecnologie di intelligenza artificiale per la salute;
- L'OMS dovrebbe supportare i paesi nella valutazione dei regimi di responsabilità che sono stati introdotti per l'uso delle tecnologie di intelligenza artificiale per la salute e in che modo tali regimi dovrebbero essere adattati ai diversi sistemi sanitari e contesti nazionali;
- L'OMS e le agenzie partner dovrebbero cercare di stabilire norme internazionali e standard legali per garantire la responsabilità nazionale per proteggere i pazienti da errori medici.

La sezione 9 presenta elementi di un quadro di governance per l'IA per la salute, con un occhio alla copertura sanitaria universale. La sezione analizza diversi framework di governance: governance

dei dati; controllo e condivisione dei benefici; governance del settore privato; governance del settore pubblico; considerazioni normative; il ruolo di un osservatorio politico e di una legislazione modello; e governance globale dell'IA: ¹²⁰

- I governi dovrebbero disporre di leggi e regolamenti chiari sulla protezione dei dati per l'uso dei dati sanitari e la protezione dei diritti individuali, compreso il diritto a un consenso informato significativo;
- dovrebbero istituire autorità indipendenti per la protezione dei dati con poteri e risorse adeguati per monitorare e far rispettare le norme e i regolamenti nelle leggi sulla protezione dei dati.;
- dovrebbero richiedere alle entità che cercano di utilizzare i dati sanitari di essere trasparenti circa l'ambito dell'uso previsto dei dati. Dovrebbero essere supportati i meccanismi per la supervisione comunitaria dei dati. Questi includono i collettivi di dati e l'istituzione della sovranità dei dati da parte delle comunità indigene e di altri gruppi emarginati.
- Gli hub di dati dovrebbero soddisfare i più elevati standard di consenso informato se i loro dati potrebbero essere utilizzati dal settore privato o pubblico, dovrebbero essere trasparenti nei loro accordi con le aziende e dovrebbero garantire che i risultati della collaborazione sui dati forniscano il massimo beneficio pubblico possibile. ¹²¹

Infine, l'allegato del rapporto fornisce consigli pratici per l'attuazione della guida dell'OMS per tre gruppi di parti interessate: sviluppatori di tecnologie di intelligenza artificiale, ministeri della salute e operatori sanitari. Le considerazioni sono intese come punto di partenza per discussioni specifiche del contesto, poiché i sistemi di IA dovrebbero essere progettati per riflettere la diversità dei contesti socioeconomici e sanitari. Questi sistemi dovrebbero essere accompagnati da formazione sulle competenze digitali, coinvolgimento della comunità e sensibilizzazione, in modo che i pazienti, le organizzazioni della comunità e la società civile siano "in grado di chiedere conto ai governi e alle aziende, di partecipare alla progettazione di tecnologie e regole , per sviluppare nuovi standard e

¹²⁰ Rohit Malpani, and About World Health Organization (WHO) More knowledge from World Health Organization (WHO). “*Ethics and Governance of Artificial Intelligence for Health: Who Guidance.*” The Communication Initiative Network, 2021, si veda <https://www.comminet.com/global/content/ethics-and-governance-artificial-intelligence-health-who-guidance>

¹²¹ Fasan, Marta “*Oms – Ethics Ad Governance of Artificial Intelligence for Health: Who Guidance.*” Biodiritto, 2021, <https://www.biodiritto.org/Biolaw-pedia/Docs/OMS-Ethics-ad-Governance-of-Artificial-Intelligence-for-Health-WHO-Guidance>

approcci e per chiedere e ricercare la trasparenza per soddisfare le proprie esigenze, nonché quelle delle loro comunità e dei sistemi sanitari."¹²²

In definitiva, sottolinea l'OMS, "l'attuazione di questa guida richiederà un'azione collettiva". Governi, fornitori e progettisti devono collaborare per affrontare i problemi di etica e diritti umani in ogni fase della progettazione, dello sviluppo e dell'implementazione di una tecnologia di intelligenza artificiale. Questo documento ha lo scopo di guidarli attraverso la revisione delle leggi esistenti e degli obblighi in materia di diritti umani e l'esplorazione di nuove leggi e politiche che sanciscono i principi etici.

8. Le sfide etiche e legali della intelligenza artificiale sanitaria negli USA

Durante la presidenza di Barack Obama, i rapporti del governo degli Stati Uniti sull'IA si sono occupati, tra le altre cose, delle applicazioni dell'IA per il bene pubblico e degli aspetti di equità, sicurezza e governance.

Dopo la presidenza di Donald Trump, la strategia di IA degli Stati Uniti si è spostata verso un approccio più orientato al libero mercato.

La Casa Bianca, ad esempio, ha ospitato l'IA per l'American Industry Summit nel maggio 2018. Una delle principali conclusioni del summit è che l'amministrazione Trump mirasse a rimuovere le barriere normative alle innovazioni dell'IA.

Nel luglio 2018, l'Ufficio Esecutivo del Presidente ha annunciato che la leadership americana nell'IA diventasse una delle principali aree prioritarie del bilancio dell'Amministrazione Usa¹²³.

Nel febbraio 2019, Trump ha firmato l'ordine esecutivo sul mantenimento della leadership americana nell'intelligenza artificiale per affrontare la critica secondo cui gli Stati Uniti hanno adottato un approccio hands-off all'AI in contrasto con altri paesi come la Cina¹²⁴.

¹²² Rohit Malpani, and About World Health Organization (WHO) More knowledge from World Health Organization (WHO). "Ethics and Governance of Artificial Intelligence for Health: Who Guidance." The Communication Initiative Network, 2021, <https://www.comminet.com/global/content/ethics-and-governance-artificial-intelligence-health-who-guidance>

¹²³ Executive Office of the President. Memorandum for the Heads of Executive Departments and Agencies, 2018, M-18-22, <https://www.whitehouse.gov/wp-content/uploads/2018/07/M-18-22.pdf>

¹²⁴ National Archives and Records Administration, <https://trumpwhitehouse.archives.gov/presidential-actions/executive-order-maintaining-american-leadership-artificial-intelligence/>.

Con questo ordine esecutivo, Trump ha lanciato una strategia coordinata del governo federale, vale a dire l'American AI Initiative, guidata da cinque aree chiave di enfasi: investire in AI, liberare risorse AI, stabilire standard di governance AI, costruire la forza lavoro AI, e impegno internazionale e protezione del vantaggio degli USA nell'IA¹²⁵.

Solo di recente, nel gennaio 2020, la Casa Bianca ha pubblicato un progetto di linee guida per la regolamentazione delle applicazioni dell'IA. Questo progetto contiene 10 principi che le agenzie dovrebbero prendere in considerazione nel momento in cui formulino approcci alle applicazioni dell'IA: fiducia del pubblico nell'IA, partecipazione del pubblico, integrità scientifica e qualità dell'informazione, valutazione e gestione dei rischi, benefici e costi, flessibilità, equità e non discriminazione.

Ci sono anche numerosi disegni di legge relativi all'IA che sono stati introdotti nel Congresso degli Stati Uniti dall'inaugurazione di Trump il 20 gennaio 2017, come il Self Drive Act, il Future of Artificial Intelligence Act del 2017 e l'AI Jobs Act del 2019.¹²⁶

Il Self drive Act è l'unico progetto di legge che ha approvato una camera (vale a dire, la Camera dei Rappresentanti degli Stati Uniti), e nessuna di queste leggi sono direttamente correlate agli aspetti etici e legali dell'IA nel settore sanitario.

Tuttavia, le due leggi del Future of Artificial Intelligence Act del 2017, ad esempio, prevedono che il Segretario del Commercio istituisca un comitato consultivo federale che studierà e valuterà, tra l'altro, come integrare le norme etiche nello sviluppo e nell'attuazione dell'IA o come lo sviluppo dell'IA possa incidere sui risparmi sui costi nel settore sanitario.

Vi sono anche sviluppi giuridici connessi all'IA a livello statale e locale¹²⁷. Ad esempio, lo Stato della California ha adottato all'unanimità la legislazione nell'agosto 2018 (ACR-215) approvando i 23 principi di Asilomar.¹²⁸ Le IA sono già in uso clinico negli Stati Uniti.

¹²⁵ US Government. Accelerating America's leadership in artificial intelligence, 2019 <https://www.whitehouse.gov/articles/accelerating-americas-leadership-in-artificial-intelligence>.

¹²⁶ Gerke, Sara, et al. "Ethical and Legal Challenges of Artificial Intelligence-Driven Healthcare." Artificial Intelligence in Healthcare, 2020, pp. 295–336., <https://doi.org/10.1016/b978-0-12-818438-7.00012-5>.

¹²⁷ FLI Team. AI policy—United States, 2019 <https://futureoflife.org/ai-policy-united-states>.

¹²⁸ Blasimme, Alessandro "Governare la biomedicina: l'etica della ricerca scientifica come questione pubblica" Paradigmi, 2011

I 23 Principi si suddividono in tre aree: Ricerca – Etica e valori – Problemi di scenario. I principi che interessano l'etica sono la maggior parte e riguardano privacy, democrazia, responsabilità, trasparenza, sicurezza, libertà e condivisione del benessere. I Principi sono quindi tesi sul futuro del mondo e dell'umanità.

In particolare, l'IA mostra grandi promesse nei settori della diagnostica e dell'imaging. Particolarmente importante in questo contesto è la Food and Drug Administration (FDA), un'agenzia federale del Dipartimento della salute e dei servizi umani, responsabile della protezione e della promozione della salute pubblica negli USA che si occupa del controllo e della supervisione di sicurezza alimentare, dei prodotti del tabacco, dei prodotti a base di caffeina, degli integratori alimentari, dei farmaci da prescrizione e da banco (farmaci), dei vaccini, dei biofarmaci, delle trasfusioni di sangue e dei dispositivi medici.¹²⁹

Con riferimento a quest'ultimi la FDA ha già autorizzato o approvato circa 40 dispositivi medici basati sull'IA¹³⁰. Ad esempio, nel gennaio 2017, Arterys ha ricevuto l'autorizzazione dalla FDA degli Stati Uniti per la sua piattaforma di imaging medico come prima applicazione ML da utilizzare nella pratica clinica¹³¹. È stato inizialmente autorizzato per l'analisi dell'immagine di risonanza magnetica cardiaca, ma lo strumento nel frattempo ha anche ricevuto l'autorizzazione dalla FDA per altri dispositivi sostanzialmente equivalenti¹³².

Idx-DR è il primo sistema diagnostico IA autorizzato dalla FDA che fornisce una decisione di screening autonoma senza la necessità per un essere umano di interpretare l'immagine o i risultati in aggiunta¹³³.

Nell'aprile 2018, la FDA ha consentito la commercializzazione di questo dispositivo basato sull'IA per rilevare più di un livello lieve della retinopatia diabetica negli adulti (dai 22 anni) con diagnosi di diabete¹³⁴.

¹²⁹ Wikipedia contributors. "Food and Drug Administration." *Wikipedia*, 16 Feb. 2023, en.wikipedia.org/wiki/Food_and_Drug_Administration.

¹³⁰ The Medical Futurist. FDA approvals for smart algorithms in medicine in one giant infographic, 2019 si veda <https://medicalfuturist.com/fda-approvals-for-algorithms-in-medicine>.

¹³¹ Leibowitz, C. "Arterys: Deep Learning for Medical Imaging. Demystifying" *Big Data and Machine Learning for Healthcare*, 2017 pp. 169–173.

¹³² Traoré, Mariam, and Shuichiro Yamamoto. "Healthcare CloudEcosystem Risk Analysis and Modeling: A FAIR Approach—A Case Study of Arterys™ on AWS." 2018, 7th International Congress on Advanced Applied Informatics, DOI:[10.1109/IIAI-AAI.2018.00171](https://doi.org/10.1109/IIAI-AAI.2018.00171).

¹³³ FDA permits marketing of artificial intelligence-based device to detect certain diabetes-related eye problems, 2018 si veda <https://www.fda.gov/newsevents/newsroom/pressannouncements/ucm604357.htm>.

Il medico carica le immagini delle retine del paziente su un server cloud e il software idx-DR fornisce quindi al dottore la raccomandazione di ripetere la scansione in 12 mesi o di mandare il paziente a uno specialista dell'occhio quando viene rilevata più di lieve retinopatia diabetica.

Nel maggio 2018, la FDA ha anche concesso l'autorizzazione alla commercializzazione del software OsteoDetect di Imagen per aiutare i medici a rilevare in pazienti adulti un tipo comune di frattura del polso, chiamata frattura del radio distale.

OsteoDetect utilizza tecniche ML per analizzare immagini a raggi X bidimensionali per identificare ed evidenziare questo tipo di frattura.

L'obiettivo principale della FDA è l'applicazione del Federal Food, Drug and Cosmetic Act,¹³⁵ un insieme di leggi approvate dal Congresso degli Stati Uniti nel 1938 che conferiscono l'autorità all'agenzia federale per realizzare i suoi compiti di supervisione.

La FDCA rileva per questo studio soprattutto per la definizione di dispositivo medico ossia "uno strumento, un apparecchio, un attrezzo, una macchina, un congegno, un impianto, un reagente in vitro o un altro articolo simile o correlato, compreso qualsiasi componente, parte o accessorio, che è 1. riconosciuto nel National Formulary ufficiale o nella United States Pharmacopeia, o qualsiasi integrazione ad essi; 2. destinato all'uso nella diagnosi di malattie o altre condizioni, o nella cura, mitigazione, trattamento o prevenzione delle malattie, nell'uomo o altri animali; o 3. destinato ad incidere sulla struttura o su qualsiasi funzione del corpo dell'uomo o di altri animali; e che non raggiunge gli scopi primari previsti attraverso l'azione chimica all'interno o sul corpo dell'uomo o di altri animali e che non dipende dalla metabolizzazione per il raggiungimento dei suoi scopi primari previsti."¹³⁶ Ad esempio, i dispositivi medici comprendono semplici abbassalingua, pacemaker con tecnologia a microchip e prodotti diagnostici in vitro come reagenti e kit di test.¹³⁷ Interessante è la "legge sulle cure del XXI secolo" (The 21st Century Cures Act) che è stata promulgata dall'ex presidente, Barack Obama, il 13 dicembre 2016.

¹³⁵ [Office of the Commissioner](https://www.fda.gov/regulatory-information/laws-enforced-fda/federal-food-drug-and-cosmetic-act-fdc-act). (2018, March 29). *Federal Food, Drug, and Cosmetic Act (FD&C Act)*. U.S. Food And Drug Administration. <https://www.fda.gov/regulatory-information/laws-enforced-fda/federal-food-drug-and-cosmetic-act-fdc-act>

¹³⁶ [Office of the Commissioner](https://www.fda.gov/regulatory-information/laws-enforced-fda/federal-food-drug-and-cosmetic-act-fdc-act). (2018, March 29). *Federal Food, Drug, and Cosmetic Act (FD&C Act)*. U.S. Food And Drug Administration. <https://www.fda.gov/regulatory-information/laws-enforced-fda/federal-food-drug-and-cosmetic-act-fdc-act>

¹³⁷ FDA. Is the product a medical device?, <https://www.fda.gov/medicaldevices/deviceregulationandguidance/overview/classifyyourdevice/ucm051512.htm>.

Inizialmente, alcuni esperti speravano che la FDA avrebbe iniziato a regolamentare completamente strumenti di consulenza medica¹³⁸.

La società Food and Drug Administration ha espresso il suo forte sostegno alla legge in un comunicato stampa, sottolineando che avrebbe sostenuto l'innovazione sanitaria e l'avanzamento della medicina di precisione negli Stati Uniti.¹³⁹

Il 21st Century Cure Act ha introdotto un'esenzione della FDCA per il software medico.

In particolare la sezione 201(h) della FDCA è stata modificata e contiene le seguenti quattro categorie di funzioni software che generalmente non rientrano nella definizione di dispositivo:

1. La funzione software è intesa "per il supporto amministrativo di una struttura sanitaria" (inclusa l'analisi aziendale, gli orari degli appuntamenti e i flussi di lavoro di laboratorio); 2. La funzione software è intesa "per mantenere o incoraggiare uno stile di vita sano e non è correlato alla diagnosi, alla cura, alla mitigazione, alla prevenzione o al trattamento di una malattia o condizione"; 3. La funzione software è intesa "a servire come cartelle cliniche elettroniche" (e "non è inteso per interpretare o analizzare le annotazioni del paziente"); o 4. la funzione del software è intesa "per il trasferimento, l'immagazzinamento, la conversione dei formati, o la visualizzazione della prova clinica del laboratorio o altri dati e risultati del dispositivo."

La FDA ha anche pubblicato una guida non vincolante per l'industria e il personale della Food and Drug Administration sulle modifiche alle politiche esistenti in materia di software medico derivanti dalla Legge sulle cure del XXI secolo¹⁴⁰ per fornire chiarimenti sull'interpretazione della Sezione della FDCA che si occupa della definizione di dispositivo.

La FDCA di contro ha anche sottolineato le caratteristiche che un dispositivo deve avere per essere esente dalla definizione del dispositivo, ossia una funzione software deve soddisfare contemporaneamente i seguenti quattro criteri: 1. La funzione software non è "destinato ad acquisire, elaborare o analizzare un'immagine medica o un segnale da un dispositivo diagnostico in vitro o un modello o un segnale da un sistema di acquisizione del segnale". 2. La funzione software è intesa "allo scopo di (...) visualizzare, analizzare o stampare informazioni mediche su un paziente

¹³⁸ IBM Watson for Oncology, è un sistema di supporto alle decisioni di cognitive computing che analizza i dati dei pazienti contro migliaia di casi storici e approfondimenti raccolti lavorando migliaia di ore con i medici del Memorial Sloan Kettering Cancer Center di New York e altri analisti.

¹³⁹ IBM letter of support for the 21st Century Cures Act, <https://www.ibm.com/blogs/policy/ibm-letter-support-21st-century-cures-act>, 2016.

¹⁴⁰ FDA. Changes to existing medical software policies resulting from Section 3060 of the 21st Century Cures Act, <https://www.fda.gov/media/109622/download>, 2019

o altre informazioni mediche (come studi clinici peer-reviewed e linee guida di pratica clinica)." 3. La funzione software è intesa "allo scopo di (...) sostenere o fornire raccomandazioni a un professionista sanitario in materia di prevenzione, diagnosi o trattamento di una malattia o condizione." 4. La funzione software è intesa "al fine di (...) consentire a tale professionista sanitario di esaminare in modo indipendente la base di tali raccomandazioni che tale software presenta in modo che non sia l'intento che tali professionisti sanitari si basano principalmente su una qualsiasi di tali raccomandazioni per fare un clinico diagnosi o decisione di trattamento riguardante un singolo paziente"

In particolare, la FDA chiarisce che il termine "supporto decisionale clinico" è definito in senso lato e significa funzioni software che soddisfano i primi due criteri e parte del terzo criterio cioè, destinato "allo scopo di sostenere o fornire raccomandazioni". Una funzione è esente dalla definizione del dispositivo solo quando è soddisfatto anche il quarto criterio. Quindi è decisivo determinare se la funzione software consente al "professionista sanitario di esaminare in modo indipendente la base per tali raccomandazioni che tale software presenta." La FDA chiarisce nella sua bozza di guida che "lo sviluppatore di software dovrebbe descrivere i dati sottostanti utilizzati per sviluppare l'algoritmo e dovrebbe includere descrizioni in linguaggio semplice della logica o razionale utilizzata da un algoritmo per rendere una raccomandazione. Le fonti a sostegno della raccomandazione o le fonti alla base della raccomandazione dovrebbero essere identificate e messe a disposizione dell'utente previsto (ad esempio, linee guida di pratica clinica con la data o la versione, letteratura pubblicata o informazioni che sono state comunicate dallo sviluppatore all'utente previsto) e comprensibile dall'utente previsto (ad esempio, punti di dati il cui significato è ben compreso dall'utente previsto)".

La FDA afferma anche che gli operatori sanitari si basano principalmente sulle raccomandazioni del software e quindi non sono in grado di "rivedere in modo indipendente la base di tali raccomandazioni".

Ci sono molte altre importanti iniziative della FDA tra cui la sua guida sul software come dispositivo medico SAMD¹⁴¹ e il lancio del cosiddetto Software Pre-Cert Pilot Program.

¹⁴¹ i SaMD (Software as a Medical Device (SaMD)) sono software che possono funzionare su piattaforme informatiche generiche (non mediche). Possono essere utilizzati in combinazione con altri dispositivi medici o altro hardware/software. Il SaMD, che di per sé è un dispositivo medico, è uno dei tre tipi di software relativi ai dispositivi medici. Gli altri due tipi comprendono il software che è parte integrante di un dispositivo medico e quello utilizzato nella fabbricazione o nella manutenzione di un dispositivo medico. I SaMD rappresentano una rivoluzione nella tecnologia sanitaria digitale, in grado di svolgere funzioni mediche complesse, diagnosticare condizioni, indicare trattamenti e supportare la gestione clinica. L'uso del software come

Quest'ultimo consentirà ad alcuni sviluppatori di salute digitale di diventare precertificati in base all'eccellenza nei criteri identificati (ad esempio, la sicurezza dei pazienti, la responsabilità clinica e la qualità del prodotto) e portare il loro software a rischio inferiore-dispositivi medici basati con revisione FDA più semplificata per commercializzare o nessuna revisione.¹⁴²

La FDA ha anche pubblicato un modello di lavoro che contiene suggerimenti per i componenti principali del programma pilota Pre-Cert. Anche se ci sono ancora molte domande aperte, il programma è un esperimento normativo innovativo che può tenere lezioni per i paesi peer e dovrebbe essere seguito da vicino.

In particolare, la FDA¹⁴³ ha proposto nell'aprile 2019, un quadro normativo per il commento pubblico per le modifiche al software basato su intelligenza artificiale/ apprendimento automatico (AI/ ML) come dispositivo medico (SaMD).

SaMD è un "software destinato ad essere utilizzato per uno o più scopi medici che eseguono questi scopi senza essere parte di un dispositivo medico hardware".

Il documento di discussione della FDA propone un nuovo approccio normativo per il ciclo di vita totale del prodotto per i dispositivi SaMD basati su IA/ML che consentono a tali dispositivi di adattarsi e ottimizzare le loro prestazioni in tempo reale per migliorare continuamente garantendo al contempo la loro sicurezza ed efficacia¹⁴⁴.

Gli sforzi fatti dalla FDA sono lodevoli, ma sarà essenziale per le autorità di regolamentazione statunitense concentrarsi in particolare sullo sviluppo di un processo per monitorare, gestire e identificare continuamente i rischi a causa di caratteristiche strettamente legate all'affidabilità dei sistemi AI/ ML.

Inoltre, quando vi è un sostanziale coinvolgimento umano nel processo decisionale, diventa ancora più difficile per le autorità di regolamentazione determinare gli effetti dell'aggiornamento di tali dispositivi.

Tutto questo rileva sicuramente sul modello di responsabilità.

dispositivo medico è in continua crescita. I SaMD possono essere utilizzati per molti scopi, inclusi l'imaging ed il monitoraggio di specifiche funzioni o parametri medici, aiutando l'operatore sanitario nel percorso diagnostico-terapeutico.

¹⁴² FDA. Changes to existing medical software policies resulting from Section 3060 of the 21st Century Cures Act, <https://www.fda.gov/media/109622/download> , 2019

¹⁴³ FDA. Proposed regulatory framework for modifications to artificial intelligence/ machine learning (AI/ML)-based Software as a Medical Device (SaMD), <https://www.fda.gov/media/122535/download> , 2019

Nel modello americano attualmente, sembra che i medici possano essere ritenuti responsabili anche se si sono affidati in buona fede a un algoritmo ML "black-box" perché il software basato sull'IA è considerato uno strumento sotto il controllo del professionista sanitario che prende la decisione finale; rimane il capitano della nave e quindi responsabile della sua rotta.

Ma la responsabilità dovrebbe configurarsi anche nel caso in cui la funzione software non consente al professionista sanitario di esaminare in modo indipendente il caso clinico che lo occupa?

Ancora, potremmo immaginare un futuro in cui l'uso della tecnologia basata sull'IA diventi lo standard di cura, e quindi la scelta di non utilizzare tale tecnologia sottoporrebbe il clinico a responsabilità?

Al momento, l'utilizzo di AI avanzata non sembra (ancora) essere parte dello standard di cura.

Pertanto, per evitare la responsabilità per negligenza medica, i medici possono utilizzare l'IA come strumento di conferma per assistere i processi decisionali esistenti anziché dover seguire le raccomandazioni dettate dalla stessa per paura di responsabilità¹⁴⁵.

La definizione del regime di responsabilità ottimale dipende in larga misura da quello che si pensa sia il "problema". Se si è preoccupati che l'impiego della tecnologia AIbased nello spazio clinico è associato ad un alto rischio per i pazienti di farsi male, si potrebbe voler mantenere l'attuale regime di *malpractice* medica che tenta di soddisfare entrambe le funzioni della legge illecito: 1) deterrenza e 2) risarcimento delle vittime. Al contrario, se si ritiene che nel corso dei casi, l'affidamento sull' IA promuove la salute del paziente, allora può essere un problema se i medici si dimostrano riluttanti a fare affidamento su questi algoritmi.

Questo potrebbe spingere il policymaker a un modello diverso.

Alcuni hanno proposto di ricondurre la responsabilità dei danni cagionati dall'IA alla responsabilità del prodotto contro i produttori di AI, un illecito che generalmente comporta una responsabilità rigorosa del produttore per difetti.

Tuttavia, ci sono notevoli sfide per vincere una tale affermazione nella pratica. I tribunali americani hanno esitato ad applicare o estendere le teorie di responsabilità del prodotto agli sviluppatori di software sanitario poiché tale software è attualmente considerato principalmente come uno strumento per supportare i medici a prendere la decisione finale.¹⁴⁶

¹⁴⁵ Price II WN, Gerke S, Cohen IG. "Potential liability for physicians using artificial intelligence" JAMA 2019; pp. 322

¹⁴⁶ Price II WN. "Artificial intelligence in health care: applications and legal implications." SciTech Lawyer 2017.

Oltre ai medici e ai maker, si deve anche considerare la responsabilità degli ospedali che acquistano e implementano i sistemi di IA.

Il risarcimento del danno in questo caso potrebbe essere richiesto in virtù delle teorie di negligenza aziendale e responsabilità indiretta.

Una teoria interessante per la responsabilità ospedaliera è quella in merito alle "credenziali negligenti". Proprio come quando gli ospedali possono essere responsabili se non rivedono adeguatamente le credenziali e la pratica dei medici e di altro personale che impiegano¹⁴⁷, analogamente possono avere doveri simili quando "assumono" un'IA.

Ancora un'altra possibilità potrebbe essere quella di abbinare uno scudo di responsabilità con un più rigoroso schema di pre-approvazione che immunizzi i professionisti della sanità e produttori da alcune forme di responsabilità a causa del processo di approvazione.

Come accade nell'esperienza europea anche nell'esperienza americana nel mondo dei big data è di fondamentale importanza che ci siano leggi di protezione in atto che proteggono adeguatamente la privacy delle persone, in particolare i pazienti.

Di particolare interesse è la Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule ossia la legge federale chiave per proteggere la privacy dei dati sanitari.¹⁴⁸

Tuttavia, l' HIPAA ha lacune significative. Ad esempio il Regolamento Americano non si applica alle informazioni sanitarie che supportano inferenze sulla salute come l'acquisto di un test di gravidanza su Amazon.

In particolare, gran parte delle informazioni sulla salute raccolte da giganti tecnologici come Amazon, Google, IBM, Facebook e Apple che stanno tutti investendo pesantemente nel campo dell'IA nel settore sanitario, non ricadono all'interno dell'HIPAA¹⁴⁹.

L'HIPAA non si applica neppure in caso di informazioni sanitarie generate dagli utenti. Ad esempio, un post su Facebook su una malattia non rientra nel regime dell'HIPAA¹⁵⁰.

¹⁴⁷ Bezaire J, Felton KW, Greve P, Allen D. "*Hospital negligent credentialing liability*" 2017 consultabile al <https://www.willistowerswatson.com/en-US/Insights/2017/08/hospital-negligentcredentialing-liability>

¹⁴⁸ Price II WN, Cohen IG. "*Privacy in the age of medical big data*". Nat Med, 2019 <https://doi.org/10.1038/s41591-018-0272-7>

¹⁵⁰ Cohen IG, Mello MM. "*HIPAA and protecting health information in the 21st century*" JAMA 2018 <https://doi.org/10.1001/jama.2018.5630>

Un problema diverso con HIPAA è la sua dipendenza dalla de-identificazione come strategia sulla privacy. Nell'ambito dell'HIPAA le informazioni sanitarie de-identificate possono essere condivise liberamente a fini di ricerca e commerciali.

Esso fornisce due opzioni per la de-identificazione: 1) una determinazione da parte di qualcuno con conoscenze adeguate ed esperienza con metodi e principi scientifici e statistici generalmente accettati; o 2) la rimozione di identificatori specifici dal set di dati come nomi, numeri di previdenza sociale e identificatori biometrici della persona o di parenti, membri della famiglia o datori di lavoro della persona, e nessuna conoscenza effettiva dell'entità coperta che le informazioni potrebbero essere utilizzate per identificare un individuo.¹⁵¹

Ma questo potrebbe non proteggere adeguatamente i pazienti a causa della possibilità di triangolazione dei dati, ovvero l'utilizzo di strumenti multipli di raccolta dei dati o di risorse dei dati per dare una maggiore profondità allo studio dei fenomeni da prospettive.

Il problema della triangolazione dei dati è stato anche recentemente descritto in una causa, *Dinerstein v. Google*¹⁵², in cui i querelanti hanno sostenuto che gli imputati hanno condiviso le cartelle cliniche con Google contenente informazioni sufficienti che hanno permesso al colosso di internet di ri-identificare potenzialmente i pazienti dato tutti i suoi altri dati a portata di mano.

Dinerstein v. Google è un chiaro promemoria del fatto che sia l'HIPAA che il diritto contrattuale non sono sufficienti per gestire questo tipo di violazioni della privacy. I pazienti vengono lasciati apparentemente indifesi contro la condivisione delle loro informazioni più personali senza il loro significativo consenso .

Nel 2017, Google ha annunciato una partnership con l'Università di Chicago per creare strumenti di apprendimento automatico in grado di prevedere i futuri problemi di salute dei pazienti e gli eventi medici avversi. Per raggiungere questo obiettivo, l'università ha consegnato a Google cinque anni di dati di cartelle cliniche elettroniche "non identificati".

Un paziente di UChicago, Matt Dinerstein, ha presentato diverse denunce contro il sistema sanitario e Google, inclusa la violazione del contratto basata su una presunta violazione dell'HIPAA. Ma un giudice federale dell'Illinois ha respinto le sue affermazioni , affermando che la parte istante non avesse dimostrato danni.

Il risultato fornisce un paio di importanti richiami.

¹⁵¹ Journal, Hipaa. “De-identification of Protected Health Information: 2023 Update.” *HIPAA Journal*, 2023, www.hipaajournal.com/de-identification-protected-health-information

¹⁵² *Dinerstein v. Google* 2019

In primo luogo, è chiaro che il diritto contrattuale americano non è uno strumento adeguato per proteggere la privacy del paziente.

HIPAA non crea un diritto privato di azione . Quindi, quando Dinerstein ha citato in giudizio il suo ospedale e Google, ha presentato reclami per violazione del contratto sulla base degli accordi sulla privacy che aveva firmato al momento del ricovero in ospedale.

Il giudice ha affermato che anche se questo accordo sulla privacy fosse stato violato, Dinerstein non avrebbe avuto diritto al risarcimento del danno, in quanto non aveva adeguatamente dimostrato che quest' ultimo fosse conseguenza della violazione.¹⁵³

Il parere mostra la difficoltà di stabilire danni monetari nel contesto della privacy dei dati sanitari. Senza un metodo chiaro per dimostrare i danni, è probabile che le rivendicazioni sulla privacy sulla salute basate su contratto falliscano.

Il successo delle rivendicazioni sulla privacy dei dati sanitari non dovrebbe essere forzato in questo quadro di diritto contrattuale. Un diritto di azione privato attraverso l'HIPAA potrebbe aiutare a colmare questa lacuna. In alternativa, rafforzare l'HIPAA per soddisfare le esigenze odierne, potrebbe eliminare la necessità di ricorsi al Giudice.

In secondo luogo, l'HIPAA può consentire ai sistemi sanitari di condividere informazioni reidentificabili con le aziende tecnologiche.

La difesa di Dinerstein ha sostenuto che l'ospedale UChicago ha condiviso dati medici che, sebbene privati di alcune informazioni identificative, sarebbero potuti essere facilmente identificati nuovamente da Google.

Gli avvocati hanno affermato che questi dati contenevano l'età dei pazienti e note mediche che sarebbero potute essere utilizzate per identificare nuovamente i pazienti se abbinate ai vasti database di consumatori che Google già possiede.

La Corte ha aggirato la questione della reidentificazione riconducendo la fattispecie nel novero delle esenzioni "approdo sicuro" dell'HIPAA, che consentono al sistema sanitario di condividere i dati per scopi di ricerca, a condizione che vengano rimossi identificatori specifici.

E' chiaro che la capacità di Google di reidentificare questi pazienti non è stata sufficiente per superare i criteri di anonimizzazione espliciti dell'HIPAA.

¹⁵³ Cohen, I. Glenn, and Michelle M. Mello. "Big data, big tech, and protecting patient privacy." *Jama* 2019: pp. 1141-1142.

Questo caso ci ricorda che senza ulteriori protezioni legali, i pazienti faranno fatica a controllare se i loro dati sanitari più personali sono condivisi con le più grandi aziende tecnologiche del mondo.¹⁵⁴ Per tutti questi motivi, l'HIPAA può essere considerato non particolarmente adeguato a proteggere la privacy sanitaria dei pazienti.

È auspicabile che la legge federale prenda sul serio la protezione dei dati sanitari non contemplati dall'HIPAA.

¹⁵⁴ J. Becker “*Insufficient Protections for Health Data Privacy: Lessons from Dinerstein v. Google*”, <https://blog.petrieflom.law.harvard.edu/>

CAPITOLO III

La responsabilità civile nell'utilizzo dell'I.A. in sanità

1. I principi generali della responsabilità medica.....pag. 101
2. I.A. e responsabilità medica per danni ai pazienti: prospettive giuridiche.....pag. 105
 - 2.1 I soggetti coinvolti.....pag. 113
 - 2.2 Case study: le decisioni cliniche e il nesso di causalità.....pag. 120
 - 2.3 Cenni sulla responsabilità della struttura sanitaria.....pag. 122
3. La recentissima proposta del Parlamento europeo e del Consiglio relativa all'adeguamento delle norme in materia di responsabilità civile extracontrattuale all'intelligenza artificiale (direttiva sulla responsabilità da intelligenza artificiale).....pag. 127

Conclusioni

1. principi generali della responsabilità medica

La responsabilità medica è la responsabilità che grava sul professionista che esercita un'attività sanitaria per i danni derivati alla salute psico-fisica del paziente ed imputabili ad errori e negligenza o da azioni compiute in violazione degli obblighi inerenti all'attività stessa.

Tale definizione mette in risalto in maniera lampante la centralità del delicato rapporto tra la tutela del diritto alla salute del cittadino e l'esercizio della professione medico-sanitaria in tutte le sue possibili declinazioni.

Occorre evidenziare, pertanto, come il concetto di responsabilità medica si riferisca all'azione di un sistema complesso in cui il paziente è destinatario di prestazioni mediche e sanitarie di ogni tipo, siano esse ospedaliere, terapeutiche, diagnostiche, chirurgiche o assistenziali, svolte da medici e personale con differenti qualificazioni.

Si ricade nell'ipotesi della responsabilità medica allorquando sussista un nesso causale tra la lesione alla salute psicofisica del paziente e l'agire dell'operatore sanitario.

Inoltre, il nesso di causalità e la stessa responsabilità del medico vengono parametrati o addirittura esclusi dal sussistere o meno di inefficienze e carenze della struttura sanitaria.

Ai fini di una compiuta trattazione del tema della responsabilità derivante dallo svolgimento di attività medico-sanitaria, preliminarmente occorre evidenziare come la medicina, ancor prima di essere considerata una scienza, debba essere inquadrata come una vera e propria arte.

Ciò in quanto ogni essere umano è unico e non ogni risultato o rimedio alla malattia può essere considerato univoco o generale, ma bensì statistico, non fornendo identica efficacia nel caso concreto e potendosi rivelare di volta in volta diversa la risposta alla terapia da parte del singolo paziente.

Quanto appena evidenziato consente di comprendere al meglio come, da una parte, i medici siano potenzialmente soggetti in misura maggiore rispetto ad altri professionisti a domande di risarcimento del danno da parte dei propri assistiti e, dall'altra, come possa risultare difficoltosa l'affermazione di responsabilità del medico nel caso concreto da parte del giudice e la conseguente quantificazione del danno risarcibile.

La prestazione medica contrattualmente preconstituita consiste in un'obbligazione di mezzi e non di risultato, per la quale il medico non può garantire la guarigione, impegnandosi piuttosto ad utilizzare i mezzi scientifici più idonei per il raggiungimento del risultato favorevole al paziente.

Il fallimento di una terapia, un esito infausto della malattia sono effetti che non è possibile ritenere automaticamente consequenziali alla sua attività in mancanza della prova di un nesso che causalmente o concausalmente ricollegli queste o altre nefaste circostanze ad una condotta professionale inadempiente.¹⁵⁵

In linea generale, nel nostro Ordinamento giuridico, la responsabilità medica viene inquadrata all'interno del più ampio filone della responsabilità professionale, disciplinata dal Capo II, Titolo III del Libro V del codice civile.

Ad una prima analisi del citato dettato normativo viene anzitutto in evidenza l'art. 2236 c.c., che esclude la responsabilità per "colpa lieve" del prestatore d'opera intellettuale, applicabile anche all'ambito sanitario e più volte invocata in dottrina quale clausola di salvaguardia generale.

La giurisprudenza della Corte di Cassazione, chiamata più volte a pronunciarsi sulla corretta interpretazione della disposizione in commento, ha chiarito come *"la limitazione di responsabilità professionale del medico chirurgo ai soli casi di dolo o colpa grave, ai sensi dell'art. 2236 cc, attiene esclusivamente alla perizia, per la soluzione di problemi tecnici di particolare difficoltà, con esclusione dell'imprudenza e della negligenza; infatti anche nei casi di particolare difficoltà, tale limitazione non sussiste con riferimento ai danni causati per negligenza o imprudenza, dei quali il medico risponde in ogni caso"* (Cass. Civ. Sez. III 12/3/2013 n. 6093).

Il menzionato intervento chiarificatore della Corte consente di introdurre nella trattazione la nozione di diligenza del professionista.

È noto come l'adempimento delle obbligazioni inerenti all'esercizio di un'attività professionale richieda un tipo di diligenza differente da quella del "buon padre di famiglia", quest'ultima generalmente richiesta ai consociati nel compimento di qualunque attività del quotidiano, ma bensì parametrata al tipo di attività svolta.

In tal senso, l'art. 1176 del c.c. al comma secondo parla di diligenza c.d. "qualificata".

Presupposto indefettibile di tale forma di diligenza è la già menzionata perizia, intesa quale apprendimento, comprensione e applicazione di quel complesso di regole tecniche proprie della categoria professionale di riferimento, che consentono di parametrare l'ambito del rischio consentito e, per l'effetto, l'area di liceità dell'intervento medico oltre la quale il professionista incorre in responsabilità.

¹⁵⁵ Matricardi Antonella, *La responsabilità Medica*, Altalex, 10/05/2018, disponibile al <https://www.altalex.com/documents/altalexpedia/2018/05/07/responsabilita-medica>

Nell'ambito dell'esercizio dell'attività medica occorre, inoltre, distinguere la responsabilità gravante sulla struttura sanitaria, da quella di cui è chiamato a rispondere il singolo medico che ha materialmente posto in essere la condotta colposa pregiudizievole per il paziente.

Nel nostro ordinamento, a seguito dell'emanazione della legge n. 24/2017 (c.d. legge Gelli-Bianco) gli aspetti della responsabilità medica sono stati definiti in maniera chiara e differenziata a seconda che la responsabilità per un determinato danno debba essere ascritta a coloro che operano presso una struttura sanitaria o alla struttura sanitaria stessa, sia essa pubblica o privata.

Mentre, infatti, oggi i medici rispondono a titolo di responsabilità extracontrattuale ex art. 2043 del Codice civile, le strutture sanitarie rispondono a titolo di responsabilità contrattuale ex art. 1218 del Codice civile, con tutte le conseguenze che ne derivano in termini di onere probatorio e di prescrizione, quinquennale per la responsabilità extracontrattuale e decennale per la responsabilità contrattuale.

L'obbligo di cura del sanitario, difatti, può sorgere per effetto di un contratto d'opera professionale stipulato con il paziente, nel qual caso sarà lapalissiana la natura contrattuale della responsabilità, o per effetto del rapporto di lavoro del medico con la struttura sanitaria, cui il paziente si sia rivolto. In tale ipotesi, si assiste ad una vera e propria separazione tra parte formale del contratto concluso tra il paziente e la struttura sanitaria e parte sostanziale, assunta dal medico, quale soggetto che materialmente esegue la prestazione.

Se l'azione viene promossa per responsabilità contrattuale, quindi nei confronti della struttura sanitaria o a fronte di specifico contratto col medico, il danneggiato dovrà provare l'esistenza del ricovero o del trattamento e semplicemente allegare l'inadempimento della controparte, rappresentato dalla lesione subita causalmente riconducibile, mentre dovrà essere la controparte, convenuta in giudizio, a dover dimostrare il corretto o impossibile adempimento della prestazione. Diversamente ai sensi dell'art. 2043 c.c., in ipotesi di responsabilità del medico ed in mancanza di stipulazione, l'onere della prova dell'attore riguarderà l'esistenza del danno, il nesso di causalità con la condotta e la colpa (o dolo) dell'agente.

Per offrire maggiore garanzia al paziente e rispondere ad istanze di tutela ed effettività del diritto alla salute, nel regime anteriore alla citata riforma, dottrina e giurisprudenza avevano sviluppato la teoria della c.d. responsabilità da contatto sociale qualificato.

Tale teoria consente di ricondurre la responsabilità del singolo medico, anche in assenza di un vero e proprio contratto stipulato col paziente, nell'alveo della responsabilità contrattuale.

Ciò in quanto si ritiene che, aldilà del generale principio del *neminem laedere*, pur in assenza di un contratto formalmente concluso e sottoscritto tra le parti, la relazione, o, se si preferisce, "il

contatto” che si instaura tra di esse, è idoneo, in ossequio ad una valutazione socialmente adeguata, a produrre obbligazioni in virtù dell’art. 1173 c.c.

Per cui, in virtù del contatto qualificato, il medico dovrà proteggere l’assistito evitando di danneggiarlo, in ossequio al dovere di buona fede, cui agli artt. 1175 e 1375 c.c.

In altri termini, tale teoria afferma l’esistenza di un’obbligazione che scaturisce da una fonte atipica, non già da un vero e proprio contratto, eppure produce effetti che ne determinano la riconducibilità nell’alveo del modello della responsabilità contrattuale, qualificandosi come tale il rapporto giuridico che si instaura tra i soggetti.¹⁵⁶

La Legge “Gelli-Bianco”, quanto al rapporto giuridico che si instaura tra medico-paziente, mediante la “decontrattualizzazione” del rapporto stesso ha dunque comportato maggiori oneri probatori a carico del secondo e, quindi, un minor rischio di condanna del sanitario.

Si osserva come tra le novità introdotte dalla legge in esame vi è anche la innovazione sul versante penale del fatto.

La responsabilità medica, difatti, non limita i suoi risvolti ai soli effetti civilistici, ma può comportare conseguenze anche sul piano penale.

In ambito penalistico, l’articolo 590-sexies c.p., così come riformato dalla legge Gelli-Bianco, prevede la responsabilità penale dei medici nei casi di omicidio colposo o lesioni cagionati nell’esercizio della professione sanitaria.

Tale responsabilità viene tuttavia esclusa per imperizia nel caso in cui il sanitario dimostri di essersi attenuto, nell’esecuzione della sua opera professionale, alle linee guida o alle buone prassi clinico-assistenziali.

Recita testualmente il citato articolo “*qualora l’evento si sia verificato a causa di imperizia, la punibilità è esclusa quando sono rispettate le raccomandazioni previste dalle linee guida come definite e pubblicate ai sensi di legge ovvero, in mancanza di queste, le buone pratiche clinico-assistenziali, sempre che le raccomandazioni previste dalle predette linee guida risultino adeguate alle specificità del caso concreto*”.

Tale disposizione sollecita da subito una riflessione.

Se è vero che la finalità dell’IA è quella di fornire all’essere umano una raccomandazione, lasciando sempre la discrezionalità sulla decisione finale, è doveroso chiedersi se la previsione proposta dalla macchina possa sostituire l’adesione alle linee guida specifiche come previsto dalla Legge 8 marzo 2017, n. 24 in tema di responsabilità professionale e quale sia l’effettiva libertà decisionale lasciata

¹⁵⁶ Chinè G., Fratini M., Zoppini A., “*Manuale di diritto civile*”, Molfetta (BA), Nel diritto editore, 2017, IX edizione p. 2181

al professionista quando è assistito da un sistema di IA. Poiché, secondo la Legge Gelli-Bianco, sono riconosciute come linee guida quelle elaborate da enti e istituzioni pubblici e privati nonché dalle società scientifiche e dalle associazioni tecnico-scientifiche delle professioni sanitarie iscritte in appositi elenchi, la scelta di ricondurre il comportamento del sistema di IA a questa situazione non è prevista dalla normativa e ci si chiede se possa, a questo punto, rientrare nella definizione di “buone pratiche clinico-assistenziali”.¹⁵⁷

Dal punto di vista penalistico, il nuovo Art. 590 sexies del Codice penale si prefigge l'intento di dare al medico la tranquillità di non essere automaticamente e vessatoriamente vittima di Procedimenti Penali basati sul binomio colpa lieve / colpa grave. Lo strumento per uscire dalla colpa medica pressoché automatica è stato, connettere sempre e comunque l'operato del sanitario alle “linee guida “. In buona sostanza, l'art. 590 sexies CP, rispetto al previgente decreto Balduzzi, non contempla il binomio surrettiziamente retribuzionista e deterministico della colpa lieve / colpa grave e, in secondo luogo, esclude la punibilità della “imperizia” qualora siano comunque state rispettate le linee guida o le buone pratiche clinico-assistenziali.¹⁵⁸

2. I.A. e responsabilità medica per danni ai pazienti: prospettive giuridiche

L'impiego dell'intelligenza artificiale in ambito medico costituisce, oggi, una realtà emergente e dagli scenari futuri sempre più consolidati, che induce inevitabilmente a riflessioni sul piano etico, giuridico ed ermeneutico.

L'IA può essere indubbiamente un valido strumento per assistere e guidare il professionista della salute nella prevenzione, nel ridurre l'incertezza correlata alla diagnosi della malattia, nel considerare quale trattamento terapeutico risulti più appropriato al caso concreto, nonché nella valutazione di rischi e benefici correlati alle cure.

Il ricorso a strumenti di IA nel delicato e complesso procedimento decisionale medico migliorerà l'assistenza ai pazienti nel quotidiano.

¹⁵⁷ Mauloni S., Mazzanti M., Buscemi L., “Nuove tecnologie e risvolti medico legali – la nuova era dell'intelligenza artificiale: profili medicolegali - the new era of artificial intelligence: medicolegal aspects” in Rivista Italiana di Medicina Legale (e del Diritto in campo sanitario), 2022, fasc. 3 pag. 679.

¹⁵⁸ Baiguera Altier Andrea, “La responsabilità medica colposa”, Diritto.it, 11/03/22 disponibile al <https://www.diritto.it/la-responsabilita-medica-colposa/>.

Ciò significa, inevitabilmente abbracciare l'idea della macchina dotata di intelligenza come parte integrante dell'equipe medica.

Il robot, in tale contesto, consente inoltre di automatizzare procedure e processi di routine in modo che i medici possano, nel frattempo, concentrarsi e riporre maggiore attenzione sui pazienti più complessi ed in condizioni critiche.

L'utilizzo di intelligenza artificiale in ambito sanitario comporta quindi il focalizzare l'attenzione clinica sul paziente che ne ha bisogno, ma allo stesso tempo prendersi cura di milioni di pazienti contemporaneamente, acquisendo un enorme bagaglio di informazioni in pochi minuti, che supera di gran lunga i casi che un medico potrebbe valutare nel corso dell'intera vita.

Il front-end finale di un applicativo di salute digitale con IA è rappresentato dal supporto alle decisioni cliniche (CDSS Clinical decision support system): il CDSS aiuta i medici negli snodi decisionali aderendo alle linee guida del setting clinico specifico ed eliminando le procedure inappropriate. È irragionevole aspettarsi che i medici abbiano seguito tutto quanto riportato dalla letteratura scientifica circa l'evoluzione della singola disciplina; si stima, ad esempio, che in medicina cardiovascolare ci sia una nuova pubblicazione ogni 2,7 minuti.¹⁵⁹

Gli indubbi benefici annunciati in termini di tutela del diritto alla salute del paziente, quali possono essere diagnosi precoci e somministrazione terapie sempre più precise e meno invasive, vengono conclamati dal Parlamento europeo che, con la Risoluzione del 16 febbraio 2017, evidenzia l'importanza di incentivare lo sviluppo di tali tecnologie.

Trattasi di un obiettivo la cui realizzazione è oggi di interesse primario per la comunità scientifica, in considerazione, altresì, dell'utilizzo oggi più che mai necessario, dei sistemi intelligenti nella gestione ed analisi dei big data della salute, aumentati esponenzialmente negli ultimissimi anni con le accentuate applicazioni dell'Internet of Things nel settore sanitario (IoMT).

L'intervento delle istituzioni europee in materia prende le mosse da una precisa constatazione: la sinergia originata dall'applicazione di un sistema di artificial intelligence nella robotica è in grado di stravolgere gli attuali assetti della responsabilità civile.

Più precisamente, il robot dotato di A.I. assume un rilevante margine di autonomia decisionale che si traduce inevitabilmente in imprevedibilità comportamentale: tali caratteristiche, di conseguenza, non consentono di inquadrare la fondamentale questione della responsabilità giuridica per danno cagionato da robot dotato di A.I. entro i confini delineati dalle soluzioni normative odierne – che

¹⁵⁹ Mauloni S., Mazzanti M., Buscemi L., “Nuove tecnologie e risvolti medico legali – la nuova era dell'intelligenza artificiale: profili medicolegali - the new era of artificial intelligence: medicolegal aspects” cit.

risultano essere quindi inadeguate – sia a livello codicistico (art.2043 c.c.), sia a livello comunitario (direttiva 85/374/CEE sulla responsabilità per danno da prodotti difettosi; direttiva 93/42/CEE sui dispositivi medici).¹⁶⁰

La capacità di assumere decisioni autonome, di sviluppare nuove soluzioni e di metterle in atto nell'ambiente che ci circonda, senza vincoli derivanti da qualsivoglia controllo o influenza esterna, non consente, infatti, di qualificare definitivamente il sistema dotato di intelligenza artificiale né in termini di mero strumento, né di prodotto, né tantomeno di dispositivo medico.

In tal modo l'utilizzo del robot in medicina cambia la prospettiva dell'interprete chiamato all'attribuzione di una qualificazione giuridica alla fattispecie in esame, sfuggendo alla rigida dicotomia del rapporto medico paziente cui siamo abituati.

La strong A.I. sembrerebbe quindi rivendicare una posizione autonoma e forse inedita nella regolamentazione giuridica, tant'è che, già oggi, si è assiste ad un accesissimo dibattito dottrinale circa la necessità di creare una nuova categoria giuridica avente caratteristiche specifiche ed implicazioni proprie.

Di ciò si fa portavoce il Parlamento europeo che, con la risoluzione citata, ha invitato la Commissione europea a valutare, nell'elaborazione di una proposta legislativa che disciplini organicamente la materia in esame, l'attribuzione di un vero e proprio status giuridico di "personalità" in capo ai robot dotati di A.I., implicante diritti e doveri precisi ma da definire, fra cui sicuramente non potrà mancare l'obbligo di risarcire il danno cagionato.

Il mondo scientifico si divide tra quelli che desiderano sacrificare la potenza dei modelli di IA a favore della esplicabilità e quelli che desiderano aumentare la performance dell'IA a discapito di questo aspetto.¹⁶¹

La diffusione di sistemi dotati di IA comporta, come si vedrà, l'insorgere di rischi di vario tipo in capo agli utenti che si avvalgono del loro utilizzo nell'esercizio di attività professionale, quali medici e operatori sanitari, ma anche e soprattutto per il destinatario finale della prestazione medica.

Questi rischi possono essere dovuti a difetti di progettazione, acquisizione dei dati o a limiti propri dei sistemi di IA.

¹⁶⁰ Nesrine Kraiem, *"Intelligenza artificiale in medicina: quale impatto sulla responsabilità medica?"* 4C Legal, 16/11/2020, <https://www.4clegal.com/vivi-lacademy/intelligenza-artificiale-medicina-impatto-responsabilita-medica>.

¹⁶¹ Fioretti Marta, *"Questioni giuridico-dogmatiche sull'intelligenza artificiale"* SalvisJuribus, 16/08/2021, <http://www.salvisjuribus.it/questioni-giuridico-dogmatiche-sullintelligenza-artificiale/>

Ad oggi è evidente che l'attuale disciplina codicistica presenta notevoli lacune nella regolamentazione del fenomeno in argomento. Il vuoto normativo circa i risvolti sul piano giuridico e di responsabilità derivante dai sistemi di intelligenza artificiale non consente il ricorso agli strumenti necessari a tutelare i diritti del danneggiato, del professionista e/o della struttura sanitaria coinvolti. Nel caso in cui si verifichi un danno alla persona, nella valutazione della condotta, è necessario considerare l'imprevedibilità e l'opacità del funzionamento dei sistemi di IA con problematiche relative all'individuazione del soggetto che dovrà rispondere del danno arrecato e difficoltà nell'inquadramento giuridico del tipo di responsabilità cui riferirsi.¹⁶²

Il nostro ordinamento giuridico, come evidenziato in premessa, non disciplina il tema della responsabilità civile per danni cagionati dall'utilizzo dei sistemi di IA.

L'interprete del diritto deve chiedersi, pertanto, se l'IA possa essere inquadrata *sic et simpliciter* all'interno delle categorie dogmatiche del diritto già esistenti oppure se sia più corretto interrogarsi, analizzare il fenomeno ed orientarsi verso la definizione di regole diverse, che siano appositamente studiate ed emanate al fine di regolamentare questo fenomeno che, non può non essere considerato una novità nell'ambito dell'attuale panorama giuridico.

Con l'emanazione dell'art. 7 della già menzionata Legge Gelli-Bianco, come si è visto, il legislatore italiano ha stabilito che, a fronte di un evento dannoso verificatosi in ambito medico, il professionista sanitario è chiamato a rispondere del suo operato ex art. 2043 Codice civile, mentre la struttura ospedaliera, sia essa pubblica o privata, continua a rispondere a titolo di responsabilità contrattuale ex art. 1218 Codice civile.

In un'ottica interpretativa rigida, ben si potrebbe far rientrare in tale ripartizione anche l'ipotesi di danno provocato da un sistema di IA. Di conseguenza, se il paziente dovesse subire un danno imputabile all'utilizzo di una macchina mossa direttamente da IA, di tale danno risponderà, a titolo contrattuale, la struttura sanitaria, ed in tal caso il paziente dovrà limitarsi ad allegare e dimostrare l'esistenza di un danno subito e la riconducibilità dello stesso alla prestazione fornita.

Il paziente potrà inoltre agire anche (o solamente) nei confronti del professionista sanitario che controllava la macchina, ma in tal caso per responsabilità extracontrattuale, con tutti i risvolti giuridici che ciò comporta sul piano dell'onere probatorio (e, non trascurabile, del termine prescrizione) che in tale ultimo caso sarà ulteriormente aggravato dalla dovuta dimostrazione della sussistenza della colpa del professionista.

Peraltro, occorre tenere a mente che i processi che contraddistinguono i sistemi di intelligenza artificiale sono complessi e difficilmente individuabili ex post nonché ex ante, in quanto

¹⁶² Fioretti Marta, “*Questioni giuridico-dogmatiche sull'intelligenza artificiale*” cit.

l'utilizzatore può fare affidamento sulla relativa certezza del risultato richiesto al robot ma non sarà in grado di individuare il percorso decisionale né tantomeno come la macchina abbia ottenuto quel risultato concreto.

Di conseguenza, non essendo tracciabile il percorso consequenziale seguito dal sistema di IA non sarà concretamente determinabile neanche l'esatto momento in cui si è verificato l'errore che ha causato l'evento lesivo e, per l'effetto, non essendovi modo di rinvenire il nesso di causalità tra il fatto ed il danno, non sarà dimostrabile la colpa, presupposto indefettibile ai fini del risarcimento del danno e alla configurazione stessa di ipotesi di responsabilità nell'ambito della responsabilità extracontrattuale.

In una prospettiva ermeneutica, ai fini di una corretta qualificazione giuridica della responsabilità del medico derivante dai danni cagionati dall'uso di I.A., bisogna preliminarmente prendere atto delle seguenti circostanze.

Anzitutto, i robot, ad una prima analisi del fenomeno, per quanto sofisticati e "umanizzati" siano, sembra che non possano essere considerati direttamente responsabili dei propri atti od omissioni che cagionino danni a terzi.

In secondo luogo, dal lato del professionista, le norme esistenti in materia di responsabilità coprono solo le ipotesi in cui la causa di un'azione o di un'omissione del robot possa essere ascritta ad uno specifico agente umano.

Infine, l'accertamento della responsabilità è preordinariamente subordinato al fatto che il professionista agente abbia potuto prevedere ed evitare il comportamento lesivo del robot.

Considerare il robot in un'ottica meramente strumentale condurrebbe alla responsabilizzazione del medico utilizzatore il quale, senza colpa, verrebbe chiamato a rispondere della scorretta o inesatta prestazione sanitaria "erogata" dall'apparecchiatura dotata di intelligenza artificiale.

Se da un lato ciò potrebbe essere legittimato con una letterale interpretazione del riferimento alla posizione di garanzia che la Costituzione (ai sensi del combinato disposto degli artt. 2 e 32) e, pedissequamente, la giurisprudenza attribuiscono al medico nei confronti del paziente (e di cui alla già menzionata teoria del contatto sociale qualificato) dall'altro si pone chiaramente in contrasto con l'inquadramento normativo della responsabilità medica in termini extracontrattuali, ormai definitivamente sancito dalla già menzionata legge Gelli-Bianco.

Sul punto, l'interprete del diritto, nell'interrogarsi circa la corretta qualificazione giuridica della responsabilità del medico utilizzatore della macchina dotata di IA, deve primariamente focalizzarsi sugli specifici compiti che, in tale ambito, debbono gravare sul medico e, in linea generale, sul legislatore.

Sarebbe infatti auspicabile, in una prospettiva di costante ed emergente utilizzo dell'IA in ambito sanitario, realtà che, come si è visto, è oggi più che mai attuale, una maggiore consapevolezza del medico utilizzatore e una massiccia informativa nei confronti del paziente, vero destinatario e “consumatore finale” dell'attività svolta dall'IA., circa i benefici connessi all'utilizzo della stessa ed agli inevitabili rischi connessi.

Proprio il menzionato richiamo alla terminologia della disciplina consumeristica di fonte comunitaria, normativa improntata alla tutela del contraente debole, rende auspicabile l'adozione di una regolamentazione “ad hoc” anche nel campo applicativo dell'IA. in ambito sanitario, attività nell'ambito della quale vi è un ampio divario tra le conoscenze del programmatore, del medico e del paziente, vero consumatore finale dell'attività svolta dall'IA.

In tale ottica, ed in assonanza alla contrapposizione professionista/consumatore, come si vedrà, la dottrina si spinge fino ad attribuire la qualifica di professionista, parte forte del rapporto giuridico, non già al programmatore, né tantomeno al medico, ma direttamente alla macchina dotata di intelligenza artificiale.

Parte della dottrina pone in evidenza come sia compito dell'operatore sanitario mediare questa comunicazione con il paziente, spiegando non solo l'uso della terminologia complessa propria di questa materia, ma anche i possibili benefici ed i rischi correlati al suo uso.

È importante che il paziente sia informato circa il grado di autonomia della macchina e circa la possibilità di non conoscere completamente i meccanismi con i quali è stato raggiunto un determinato risultato. Per questo motivo, la trasparenza dei sistemi della IA deve essere al massimo livello al fine di aumentare la fiducia del professionista nei confronti del sistema e, di conseguenza, la fiducia dei pazienti nei confronti del professionista sanitario, necessari per un processo decisionale condiviso.¹⁶³

Una tale conclusione, però, non tiene conto dell'impossibilità, da parte del medico utilizzatore del robot, di comprendere a pieno i meccanismi con i quali i sistemi di IA raggiungono un determinato risultato.

Ciò può, viceversa, indurre il professionista sanitario ad affidarsi ciecamente alla macchina, con il rischio di perseguire condotte in grado di arrecare un danno ai pazienti, senza essere in grado di giustificare il suo operato e rischiando, paradossalmente, di incorrere in ipotesi di responsabilità.

¹⁶³ Mauloni S., Mazzanti M., Buscemi L., “Nuove tecnologie e risvolti medico legali – la nuova era dell'intelligenza artificiale: profili medicolegali - the new era of artificial intelligence: medicolegal aspects” cit.

Si pensi, in tal senso, al diritto al consenso informato, ormai riconosciuto dalla giurisprudenza della Corte di Cassazione quale vero e proprio corollario del diritto alla salute sancito all'articolo 32 comma 1 della Costituzione e del diritto all'autodeterminazione in ambito sanitario di cui al comma 2, il quale recita testualmente che *“nessuno può essere obbligato a un determinato trattamento sanitario se non per disposizione di legge.”*

Secondo la giurisprudenza di merito *“in tema di attività medico-chirurgica, il consenso informato deve basarsi su informazioni dettagliate, idonee a fornire al paziente la piena conoscenza della natura, della portata e dell'estensione dell'intervento medico-chirurgico, dei suoi rischi, dei risultati conseguibili e delle possibili conseguenze negative, non essendo all'uopo idonea la sottoscrizione da parte del paziente di un modulo del tutto generico”*. (Corte appello Torino sez. III, 09/03/2022, n.284)

Ancora, la suprema Corte ritiene che *“in tema di responsabilità civile derivante dall'esercizio professionale dell'attività medico-chirurgica, a fronte dell'allegazione del relativo inadempimento da parte del paziente, il medico è tenuto a provare di avere adempiuto all'obbligazione di avergli fornito un'informazione completa ed effettiva sul trattamento sanitario e sulle sue conseguenze.”* (Sez. 6 - 3, Ordinanza n. 27279 del 16/09/2022)

Il consenso, qualora non sia libero, ossia prestato da paziente che non abbia piena consapevolezza o addirittura non prestato in assenza di condizioni che legittimino tale circostanza, può esporre di per sé il medico o la struttura che non lo hanno acquisito a responsabilità giuridica.

Queste evidentissime ed attualmente inevitabili asimmetrie informative vengono scontate a tutti i livelli.

In tal senso, appare ancor più evidente la necessità di idonea e apposita tutela giuridica dei diritti alla salute e all'autodeterminazione ma, di converso, anche della posizione del professionista, sia esso il singolo medico o la struttura sanitaria, che non dovrebbero essere chiamati a rispondere della mancanza di dettagliate conoscenze che esulano dalle proprie competenze in ordine al pieno rispetto del diritto al consenso informato del paziente, in un ambito così tecnico e complesso, né di un errore della macchina che, seppur statisticamente più improbabile dell'errore umano, è comunque possibile.

Viceversa, qualora il legislatore rimanesse inerte al cospetto della sempre più diffusa applicazione di nuove tecnologie di prognosi della malattia, cura e assistenza del paziente, si andrebbe incontro, per via giurisprudenziale, alla possibile configurazione di una responsabilità oggettiva dell'operatore sanitario o della struttura, come già sta avvenendo anche in seno alla Comunità europea.

Il timore dettato dalla formulazione di un'ipotesi di responsabilità oggettiva, peraltro, potrebbe determinare l'impiego sempre più sporadico di robot dotati di A.I. da parte della classe medica, portando alla messa a punto di una nuova species del noto fenomeno di c.d. medicina difensiva.

Posto l'obbligo del medico non solo di attivarsi a prestare le cure al paziente ma anche di impedire l'insorgenza o l'aggravamento di condizioni patologiche che siano prevedibili ed evitabili, l'attuale quadro giuridico risulta, in definitiva, del tutto insufficiente a garantire al paziente il fondamentale diritto al risarcimento dei danni cagionati dalla nuova generazione di robot, dotati di capacità di adattamento e di apprendimento tali da implicare un significativo e preoccupante grado di imprevedibilità.¹⁶⁴

Da un punto di vista civilistico, come visto, la strada per la configurazione della responsabilità medico-sanitaria viene tracciata dal binomio responsabilità contrattuale, di maggior tutela del paziente, o extracontrattuale, volta quindi alla salvaguardia degli interessi del professionista.

Ma a ben vedere soluzioni diverse sono possibili.

Occorrerà, in tal senso, riconsiderare la visione non più così utopica di coloro che ipotizzato di risolvere il problema con una "salomonica" ripartizione del danno tra tutti i protagonisti, anche attraverso una predefinita quantificazione tabellare, risarcibile attraverso l'istituzione obbligatoria di un fondo ad hoc o garantita da forme assicurative obbligatorie. (Chi paga per l' algoritmo: le questioni ancora aperte su IA e responsabilità medica)

O, ancora, abbracciare ed "accettare" l'idea offerta in prospettiva futura dal legislatore comunitario, e cioè l'attribuzione di una vera e propria personalità giuridica ai sistemi di I.A.

Portando tale teoria ai suoi estremi teorici l'interprete del diritto assisterebbe, in tal caso, ad una poderosa rivoluzione copernicana nell'ambito della responsabilità medico-sanitaria che, attraverso il capovolgimento del centro d'imputazione della responsabilità, renderebbe protagonista, anche in ambito giuridico, non più l'uomo ma bensì la macchina.

Occorre infatti evidenziare come l'ampia diffusione di macchine dotate di IA comporta inevitabilmente la comparsa di rischi in capo all'utente finale, rischi non direttamente imputabili al singolo in sé considerato, sia esso il programmatore o il medico.

Questi rischi sono legati in parte legati ai limiti dell'intelligenza umana, responsabile delle fasi di progettazione o acquisizione dei dati, ed in parte a limiti e, si noti, sviluppi inattesi, propri dei sistemi di IA.

L'uso di sistemi di IA può generare discriminazioni legate a sesso, età, etnia, religione o in ragione della loro disponibilità ed accesso nei diversi Paesi e solleva perplessità circa l'ampia condivisione di una mole ingente di dati sensibili con minaccia ai diritti fondamentali dell'uomo.

¹⁶⁴ Nesrine Kraiem, *"Intelligenza artificiale in medicina: quale impatto sulla responsabilità medica?"* cit.

Nonostante gli sforzi posti in essere per rendere l'IA “affidabile” e “spiegabile” è evidente la difficoltà concreta nell'adottare misure tecniche e organizzative volte ad assicurare il controllo su privacy ed autonomia: i concetti di trasparenza, riservatezza, selezione e minimizzazione dei dati, oggetto delle normative attualmente vigenti, devono trovare piena applicabilità nonostante le difficoltà correlate alla complessità dell'IA e alla sua costante evoluzione.

Non manca chi sostiene che la capacità autonoma dei sistemi di IA di adottare decisioni, di metterle in atto e di auto-apprendere e innovarsi è una concatenazione che interrompe, palesemente, il nesso causale tra la condotta umana – del produttore, del programmatore della macchina, del medico o della struttura sanitaria – e del danno eventualmente causato da tale tecnologia dotata di intelligenza propria.

Pertanto, anche l'I.A. sembra legittimata ad esigere nuovi spazi di regolamentazione giuridica e, premesse le già menzionate considerazioni, appare opportuno riflettere sulla introduzione di una nuova forma di responsabilità prevista ad hoc per la sola macchina dotata di intelligenza artificiale. Si tratterebbe di una attribuzione che porterebbe alla previsione di una nuova e correlativa forma di responsabilità, esattamente come si è avuto per gli enti con il D.lgs 231/2001.

(Chi paga per l'algoritmo: le questioni ancora aperte su IA e responsabilità medica)

Ciò non significa, in ogni caso, esonerare del tutto il medico da eventuali responsabilità correlate all'esercizio della professione, bensì ripartire tali responsabilità esattamente come avviene per la divisione delle attività da compiere nella somministrazione della cura al paziente, in un'ottica di collaborazione sinergica tra essere umano e robot intelligente, tenendo sempre a mente che *“nessun robot, nessun algoritmo potrà mai sostituire il medico, perché l'arte professionale sta nell'interpretare i dati secondo le esigenze del paziente”*.¹⁶⁵

2.1 I soggetti coinvolti

Analizzate in prospettiva le possibili evoluzioni della responsabilità medica a fronte dell'utilizzo di macchine dotate di IA, ed in particolare l'aspetto della possibile configurazione di personalità giuridica in capo agli stessi robot intelligenti e la conseguente possibilità di imputazione diretta del

¹⁶⁵ Così Filippo Anelli, Presidente della FNOMCeO, in occasione degli Stati generali della professione medica del 16 e 17 maggio 2019, Roma.

danno da essi cagionato, occorre ancora una volta rilevare come, al momento attuale, l'ordinamento giuridico nazionale e comunitario non abbiano positivizzato alcuna normativa in tal senso.

Con riguardo, in particolare, al tema della responsabilità civile per i danni causati dai robot lo stesso

Parlamento europeo nella Risoluzione del 16 febbraio 2017, ha già evidenziato che «almeno nella fase attuale, la responsabilità deve essere imputata a un essere umano e non a un robot».

Nella prospettiva delineata la riparazione del danno prodotto viene dunque ancora affidata, nella tendenzialmente comune rappresentazione dei sistemi intelligenti in termini di strumenti, attraverso quel modello di responsabilità che muove dalla produzione del bene, ma può non escludere, in relazione allo specifico caso concreto, la posizione dell'utilizzatore-proprietario.¹⁶⁶

Premesso che in base ai principi generali del diritto la responsabilità per il danno ricade su chi crea, espone o mantiene un rischio e che lo stesso individuo è tenuto a minimizzarlo ex ante o risarcirlo ex post, da ciò discende che, tra i soggetti passibili di responsabilità, nel caso che ci occupa, possa rientrare: chi crea il sistema di IA, chi si occupa della manutenzione, chi lo usa nonché qualunque agente esterno che interferisca con la macchina.

Occorre quindi interrogarsi e analizzare le diverse figure coinvolte nella creazione ed uso dei robot dotati di I.A. e il regime di responsabilità applicabile ai vari soggetti secondo i comuni canoni civilistici.

Tra i diversi soggetti che concorrono all'applicazione dell'I.A. in medicina si va da chi adotta un sistema autonomo all'interno della propria organizzazione di mezzi e risorse (nella gran parte dei casi la Struttura sanitaria) a chi se ne avvale materialmente (il medico) e a tutti gli attori coinvolti nella sua realizzazione (il produttore dell'hardware, l'ideatore dell'algoritmo, il programmatore del software, il trainer, il soggetto addetto alla manutenzione del bene durante il suo ciclo di vita). Secondo il già consolidato orientamento giurisprudenziale oggi fatto proprio dalla già vista legge in materia (L. Gelli Bianco del 2017) il rapporto che lega la struttura sanitaria al paziente ha fonte in un contratto c.d. di "spedalità" o di "assistenza sanitaria", ossia il contratto in forza del quale la struttura sanitaria si obbliga a fornire al paziente una complessa prestazione di assistenza sanitaria, comprensiva di attrezzature sanitarie e personale medico sufficiente ed efficiente.

¹⁶⁶ Dr. Franco Abbenda, *"I sistemi di intelligenza artificiale come strumento di supporto alla diagnostica"* nota, Ministero della Salute, Consiglio Superiore di Sanità Sessione LII (2019-2022) Presidente: Prof. Franco Locatelli Sezione V* Presidente: Prof. Giuseppe Remuzzi Segretario tecnico: Dr. Franco Abbenda, https://www.salute.gov.it/imgs/C_17_pubblicazioni_3218_allegato.pdf

La responsabilità della struttura sanitaria è dunque di tipo contrattuale e sussiste ex art. 1218 del Codice civile, ovvero ex art. 1228 “per fatto degli ausiliari” e come tale deve intendersi il medico che esegue la prestazione.

A norma dell’art 7 della citata legge, difatti, “la struttura sanitaria pubblica o privata che, nell’adempimento della propria obbligazione, si avvalga dell’opera di esercenti la professione sanitaria, anche se scelti dal paziente e ancorché non dipendenti della struttura stessa, risponde, ai sensi degli articoli 1218 e 1228 c.c., delle loro condotte dolose o colpose”.

Quanto alla ripartizione dell’onere probatorio “in tema di responsabilità contrattuale della struttura sanitaria, incombe sul paziente che agisce per il risarcimento del danno l’onere di provare il nesso di causalità tra l’aggravamento della patologia (o l’insorgenza di una nuova malattia) e l’azione o l’omissione dei sanitari, mentre, ove il danneggiato abbia assolto a tale onere, spetta alla struttura dimostrare l’impossibilità della prestazione derivante da causa non imputabile, provando che l’inesatto adempimento è stato determinato da un impedimento imprevedibile ed inevitabile con l’ordinaria diligenza.”(Sez. 3 - , Ordinanza n. 26700 del 23/10/2018)

Per ciò che concerne la responsabilità del singolo medico, abbiamo già visto come lo stesso possa essere chiamato a rispondere personalmente nei confronti del paziente ma, secondo la normativa vigente, ex art. 2043 c.c.

Ciò comporta in capo al paziente che agisce in giudizio l’assolvimento di un diverso e più pregnante onere probatorio, in quanto sarà tenuto a dimostrare quantomeno la colpa dell’agente, nelle forme della negligenza, imprudenza e imperizia.

Per la concorde dottrina, la negligenza sussiste nei casi di noncuranza, da intendersi quale difetto di attenzione, mentre l’imprudenza si verifica nei casi di di avventatezza e/o di insufficiente ponderazione dell’agente.

L’imperizia, invece, contrariamente alla "perizia", attiene a ciò che un medico qualificato all’esercizio della professione dovrebbe essere in grado di fare secondo un modello comune di riferimento. In sostanza, dunque, l’imperizia si caratterizza per l’inosservanza, nel caso concreto, della c.d. "leges artis", ossia per avere il medico violato una regola specialistica e/o tecnica di comune conoscenza per la categoria.

Anche nell’uso dell’I.A., e anche allorquando il danno sia materialmente provocato dalla macchina, ad es. nell’ambito di un intervento chirurgico realizzato a distanza, il medico dovrà in ogni caso avere un grado di attenzione elevata, che riguardi la scelta e l’opportunità del tipo di intervento in riferimento alla situazione clinica del paziente, la scelta della macchina da utilizzare etc...

Viceversa, in ambito diagnostico, anche a fronte di un errore della macchina, il medico debitamente formato dovrebbe essere comunque in grado (nella maggior parte dei casi) di riscontrare l'erroneità dell'output inviato dal sistema di I.A. e di non "lasciarsi ingannare" nella diagnosi di una malattia, a fronte del sussistere dei classici elementi sintomatici immediatamente percepibili.

In tal senso, il medico non potrà (sempre) andare esente da responsabilità invocando l'erroneità del dato fornito dal robot ed insita in eventuali difetti di programmazione (o in imprevedibili ed autonomi sviluppi della macchina).

Tutto ciò in quanto, non va dimenticato, l'arte medica ha fatto a meno dell'IA in particolare per migliaia di anni.

In ogni caso, l'eventuale colpa del medico, nelle forme su esplicitate, non esonera il paziente dalla prova del nesso eziologico, sotto il duplice profilo della causalità materiale, ossia la riconducibilità dell'evento lesivo alla condotta realizzata dal sanitario, sia sotto il profilo della causalità giuridica, intesa quale individuazione delle singole conseguenze pregiudizievoli derivanti dall'evento, prova che potrà essere raggiunta anche mediante presunzioni.

In altri termini, la semplice violazione delle regole di diligenza, la mancanza di perizia, non è, di per sé, causa della produzione del danno- evento, "poiché l'aggravamento della situazione patologica o l'insorgenza di nuove patologie non sono immanenti alla violazione della *leges dell'ars medica* e potrebbero avere una diversa eziologia. Pertanto, la mera allegazione della causalità materiale non assorbe ipso iure la causalità giuridica; allegare il solo inadempimento non significa allegare il danno-evento, che ha ad oggetto un interesse ulteriore che corrisponde alla lesione dell'interesse presupposto, in termini di danno alla salute." ¹⁶⁷

La riconducibilità del rapporto nell'alveo della responsabilità extracontrattuale comporta inoltre l'applicabilità di un diverso regime prescrizioneale, che è quinquennale e non decennale come si verifica invece nel rapporto di assistenza sanitaria intercorrente con la casa di cura.

Come è facile notare, a questo punto, la "decontrattualizzazione" del rapporto medico-paziente operata dalla Gelli Bianco ed il conseguente abbandono, in materia, della teoria del "contatto sociale qualificato" ha comportato maggiori oneri probatori per il paziente danneggiato dalla prestazione ed un minor rischio di condanna per il medico.

Quanto alla responsabilità degli altri soggetti coinvolti nel processo produttivo e applicativo del robot, appare chiaro come nel caso di errore insito nella programmazione del software si dovrà

¹⁶⁷ Ascolese Raffaella, "*La natura giuridica della responsabilità medica: il panorama legislativo dopo la Legge n. 24 del 2017*" D.it, 21/03/2022, <https://www.diritto.it/la-natura-giudica-della-responsabilita-medica-il-panorama-legislativo-dopo-la-legge-n-24-del-2017/>

indagare circa la responsabilità del venditore o del programmatore, così come previsto dalla direttiva 85/374/CE.

Ed infatti, tanto l'algoritmo quanto i dati d'addestramento del sistema intelligente svolgono una funzione preponderante nell'economia di funzionamento del prodotto così che pare ragionevole predicare una loro responsabilità che vada oltre i rapporti contrattuali con il produttore del bene finale e si estenda anche nei confronti dei terzi danneggiati (Nicci, Vetrugno, *Intelligenza artificiale e responsabilità nel settore sanitario*, in *Intelligenza artificiale – il diritto, i diritti e l'etica*, a cura di Ruffolo, Milano, 2020, 614.)

Andrà in tal caso applicata la normativa in materia di responsabilità da prodotto difettoso, che, quanto al nostro ordinamento, è contenuta negli artt. 114-127 del Codice del consumo.

Secondo il Codice il produttore è ritenuto responsabile quando i danni causati da malfunzionamenti del prodotto siano eziologicamente riconducibili ad un "difetto" del prodotto stesso.

Ma, atteso che, secondo la tesi maggioritaria, la responsabilità da prodotto difettoso non ha natura oggettiva¹⁶⁸ va da subito evidenziato che l'onere che grava sul danneggiato di dimostrare il difetto del prodotto e il nesso di causalità tra difetto e danno rischia di diventare probatio diabolica proprio in considerazione di quell'opacità delle modalità di funzionamento del mezzo, che non conduce ad una facile individuazione del difetto del prodotto, né all'individuazione del momento causativo del danno, e dunque, di conseguenza, del nesso di causa di cui è necessaria la prova a fini risarcitori.¹⁶⁹

Altra soluzione proviene dagli Stati Uniti, dove nell'ambito del danno provocato dal difetto dell'algoritmo si è diffusa la prassi di ricorrere al cd. risk utility test, che consiste nel verificare se il danno si sarebbe comunque prodotto anche se fosse stato utilizzato un algoritmo alternativo rispetto a quello che lo ha generato, test che, come ovvio, richiede la consulenza di personale specializzato, e dunque è di non facile accesso economico.¹⁷⁰

Inoltre, occorre sottolineare che nell'Unione Europea l'intelligenza artificiale non si è ancora sviluppata come negli Stati Uniti. Tuttavia, quello americano non è certo un modello da seguire, essendo questo caratterizzato da un deregolamento che lascia spazio a vuoti di tutela e ad abusi.

¹⁶⁸ Votano Giovanni, "*Intelligenza artificiale in ambito sanitario: il problema della responsabilità civile*", in *Danno e Responsabilità*, 2022, n. 6 p. 669.

¹⁶⁹ Fusaro Arianna, "*Quale modello di responsabilità per la robotica avanzata? Riflessioni a margine del percorso europeo*" in *La nuova giurisprudenza civile commentata*, 2020.

¹⁷⁰ Coletti Elisa, "*Intelligenza artificiale e attività sanitaria. profili giuridici dell'utilizzo della robotica in medicina*" in *Rivista di Diritto dell'Economia, dei Trasporti e dell'Ambiente*, 2021, Vol. XIX, pp. 201-219.

Diversamente, le autorità europee scelgono di normare per primi, esattamente come è accaduto per il trattamento dei dati personali.¹⁷¹

A questo proposito la proposta di Regolamento UE n. A9- 178/2020 ha prospettato la possibilità di applicare il regime di responsabilità oggettiva del produttore per i sistemi che vengono definiti “ad alto rischio”.

Esempio di responsabilità oggettiva prevista dall’ordinamento italiano in tale ambito e applicabile al caso che ci occupa è dato dall’ art. 2050 c.c. rubricato «Responsabilità per l’esercizio di attività pericolose» il quale troverebbe applicazione se l’IA, per un difetto intrinseco, dovesse essere ritenuta ex post pericolosa già nel momento in cui è stata messa in circolazione.

E ciò in ragione del cosiddetto rischio di sviluppo gravante sul produttore, il quale, sebbene privo, al momento dell’immissione in commercio, delle conoscenze tecniche e scientifiche necessarie per riconoscere il difetto, avrebbe comunque l’onere di adottare opportune cautele e adeguate precauzioni.

Parte della dottrina ha inoltre paventato l’applicabilità dell’art 2047 c.c. a norma del quale “*in caso di danno cagionato da persona incapace di intendere o di volere, il risarcimento è dovuto da chi è tenuto alla sorveglianza dell’incapace, salvo che provi di non aver potuto impedire il fatto.*”

Si potrebbe, forse, in questo caso individuare un’analogia tra il proprietario del dispositivo dotato di intelligenza artificiale, specialmente se caratterizzato quest’ultimo da un elevato grado di autonomia, e colui che è giuridicamente responsabile del soggetto di diritto che compie l’azione dannosa e che non è in grado di ricevere su di sé gli effetti giuridici della medesima, perché minore, o incapace.¹⁷²

Con riferimento a ipotesi di responsabilità oggettiva, c’è chi ha ravvisato un eccessivo squilibrio di oneri e rischi in capo a chi è chiamato a rispondere per i danni prodotti da IA, con conseguente possibili incremento delle cause di risarcimento, evidenza che “in tal modo il produttore sarebbe sempre responsabile per i danni che vengono cagionati dalla macchina, indipendentemente dalla presenza di eventuali personali profili soggettivi di dolo o di colpa”.¹⁷³

La Commissione europea, quasi a voler mediare tra le differenti istanze, ha evidenziato che in futuro “oltre alla valutazione del rischio effettuata prima dell’immissione sul mercato del prodotto, potrebbe essere introdotta un’ulteriore procedura di valutazione del rischio, da effettuarsi nel caso

¹⁷¹ Fioretti Marta, “*Questioni giuridico-dogmatiche sull’intelligenza artificiale*” cit.

¹⁷² Coletti E., “*Intelligenza artificiale e attività sanitaria. profili giuridici dell’utilizzo della robotica in medicina*” cit.

¹⁷³ Fioretti M., “*Questioni giuridico-dogmatiche sull’intelligenza artificiale*” cit.

in cui il prodotto subisse importanti modifiche durante il ciclo di vita, ad esempio una funzione diversa, non prevista dal fabbricante nella valutazione iniziale del rischio. Detta valutazione dovrebbe riguardare l'impatto sulla sicurezza dovuto al comportamento autonomo durante l'intero ciclo di vita del prodotto".¹⁷⁴

Tale ultima soluzione proposta dalla Commissione potrebbe essere dirimente e risolutiva in ordine alla problematica applicazione di ipotesi di responsabilità oggettiva gravante sul produttore di beni che, si evolvono autonomamente nel corso del tempo.

Ulteriore protagonista nella materia che ci occupa, figura particolare e fattore determinante nella fase di apprendimento del sistema di I.A., è il trainer (c.d. addestratore).

Trattasi di un soggetto che assume un ruolo essenziale ai fini del buon funzionamento della macchina e deputato allo svolgimento dell'attività di training data e di testing del sistema. La prima consiste in un'attività di raccolta, analisi e selezione dei dati da fornire all'algoritmo e sulla base dei quali quest'ultimo si "eserciterà" alla ricerca di soluzioni al problema che gli viene sottoposto. La seconda, come facilmente intuibile, in attività di controllo circa l'esattezza e l'appropriatezza delle soluzioni offerte. Il trainer può anche non coincidere col produttore del bene finale o con il programmatore.

In riferimento alla responsabilità del trainer, nonostante la già evidenziata importanza dell'attività svolta ai fini del corretto funzionamento della macchina, occorre rilevare anzitutto come il lavoro di questo soggetto, nonché il rapporto obbligatorio instaurato col produttore del bene finale, si esauriscano prima della commercializzazione del prodotto finale.

Ne consegue che alla figura in esame, in caso di comportamenti devianti nella fase successiva, non potrà essere applicata la disciplina prevista per la responsabilità da prodotto difettoso. Il terzo danneggiato dal sistema, dunque, non potrebbe percorrere la strada dell'azione diretta nei confronti del trainer, bensì potrebbe agire esclusivamente in base alla disciplina in tema di responsabilità extracontrattuale.

Dalla narrativa che precede ben si evince come, nel vigore della Legge Gelli Bianco ed in un'ottica di convenienza processuale dell'azione giudiziaria da intraprendere, il paziente danneggiato dall'uso di robot intelligenti verrà indirizzato dal professionista avvocato verso l'esercizio

¹⁷⁴ Si veda: la relazione della commissione al parlamento europeo, al consiglio e al comitato economico e sociale europeo, Relazione sulle implicazioni dell'intelligenza artificiale, dell'Internet delle cose e della robotica in materia di sicurezza e di responsabilità, <https://op.europa.eu/it/publication-detail/-/publication/4ce205b8-53d2-11ea-aece-01aa75ed71a1/language-it/format-PDF>.

dell'azione risarcitoria esclusivamente nei confronti della struttura sanitaria nell'ambito della quale l'evento lesivo si è verificato, chiamata a rispondere ex artt. 1218 e 1228 del Codice civile.

2.2 Case study: le decisioni cliniche e il nesso di causalità

Come noto, la Legge 2017, n. 24, facendo espresso riferimento alla disciplina contenuta nell'art. 2043 c.c., prevede che il paziente danneggiato sia tenuto a fornire idonea dimostrazione del nesso di causalità tra la condotta del medico e il danno da risarcire.

Tale onere può essere assolto anche per presunzioni, e cioè dimostrando che la condotta del sanitario è stata, secondo il criterio del "più probabile che non", la causa del danno.

Dal punto di vista della responsabilità civile, il problema del nesso di causa riguarda essenzialmente due profili: l'ampiezza della responsabilità del soggetto convenuto in giudizio a seguito di una sua condotta illecita e la selezione delle eventuali conseguenze dannose derivanti dalla sua condotta attiva od omissiva di cui può essere chiamato a rispondere.¹⁷⁵

Tanto premesso, occorre ancora una volta rilevare che l'introduzione di strumenti di Intelligenza Artificiale nell'erogazione della prestazione sanitaria abbia già sollevato in dottrina ampie perplessità quanto alle concrete modalità di accertamento del nesso eziologico che deve intercorrere tra l'azione (o l'omissione) del medico (che si avvale di tale tecnologia nell'esercizio della professione) e l'evento lesivo.

In particolare, diversi autori evidenziano come l'autonomia e, in un certo senso, l'imprevedibilità che caratterizzano le tecnologie in parola implicino, necessariamente, un limite alla verificabilità del nesso eziologico che deve sussistere tra la condotta dell'agente e l'evento dannoso, incidendo inevitabilmente sulla comune separazione tra condotta umana ed artificiale.

Ed infatti, poiché gli algoritmi di queste macchine sono suscettibili di autonomi cambiamenti che originano direttamente dall'esperienza, esse hanno comportamenti che non possono essere sempre preveduti ex ante dall'essere umano.¹⁷⁶

In tal senso, il comportamento autonomo del sistema, dotato di intelligenza artificiale e per questo capace di sviluppare ulteriormente proprie e ulteriori soluzioni ai problemi e quesiti che gli

¹⁷⁵ Monateri Pier Giuseppe e Gianti Davide, "*Nesso di causalità (dir. civile)*" Diritto online, 2016, [https://www.treccani.it/enciclopedia/nesso-causale-dir-civ_\(Diritto-on-line\)/](https://www.treccani.it/enciclopedia/nesso-causale-dir-civ_(Diritto-on-line)/)

¹⁷⁶ Fioretti Marta, "Questioni giuridico-dogmatiche sull'intelligenza artificiale" cit.

vengono sottoposti, configurerebbe un elemento aggiunto e come tale di per sé interruttivo del nesso di causalità. Una circostanza nuova, ed estranea alla fattispecie, in grado di introdurre un rischio notevolmente maggiore rispetto a quello che il medico abitualmente deve scontare nell'esercizio della professione sanitaria.

Ma, a ben vedere, così non è.

Orbene, una pacifica soluzione del quesito che ci occupa, allo stato attuale ed in mancanza di apposita disciplina in materia, sottintende necessariamente un'analisi della disciplina generale prevista e applicata nel nostro ordinamento in tema di rapporto di causalità e del c.d. "evento interruttivo".

Una compiuta normativa del tema oggetto d'indagine è rinvenibile in ambito penalistico, nell'elaborazione del concetto di causalità, difatti, dottrina e giurisprudenza sono stati largamente influenzati dalla normativa penalistica.

Orbene, l'art. 41 c.p. al primo comma sancisce che *"il concorso di cause preesistenti o simultanee o sopravvenute, anche se indipendenti dall'azione od omissione del colpevole, non esclude il rapporto di causalità fra l'azione od omissione e l'evento"*.

L'ipotesi che maggiormente rileva ai fini della trattazione è contenuta al secondo comma dello stesso articolo, laddove il Codice chiarisce che *"le cause sopravvenute escludono il rapporto di causalità quando sono state da sole sufficienti a determinare l'evento"*.

A ben vedere però anche i riferimenti normativi di matrice penale nulla chiariscono circa la nozione di causalità in sé considerata e riguardo ai criteri da applicare per l'accertamento della medesima, tanto che l'intera elaborazione in materia ha origine dottrinale e giurisprudenziale in quanto demandata agli interpreti.

Secondo la giurisprudenza della Suprema corte *"è configurabile l'interruzione del nesso causale tra condotta ed evento quando la causa sopravvenuta innesca un rischio nuovo e incommensurabile, del tutto incongruo rispetto al rischio originario attivato dalla prima condotta"*. (Cass. pen. n. 25689/2016)

Occorrerà dunque interrogarsi di volta in volta se, nella concausazione dell'evento lesivo tra condotta del sanitario e fatto proprio del robot, l'azione di quest'ultimo possa considerarsi di per sé interruttiva del rapporto di causalità che lega la condotta del primo all'evento e quindi in grado, autonomamente e a prescindere dalla perizia del medico, di scatenare l'evento e produrre il danno al paziente.

Proprio il richiamo alla perizia e a quanto evidenziato nei paragrafi che precedono consente di individuare la fattispecie che probabilmente, avrà più larga diffusione in ambito giuridico ogniqualvolta il medico e la struttura sanitaria siano chiamate a rispondere del danno prodotto da

IA, e cioè la mancata vigilanza, il non essersi accorti dell'errore, il non avere impedito il verificarsi del danno o, secondo la terminologia più frequente, il c.d. illecito omissivo.

L'accertamento del collegamento causale tra illecito omissivo ed evento dannoso nel settore della responsabilità medica da sempre alimenta un ampio ed acceso dibattito dottrinale e giurisprudenziale, atteggiandosi a "laboratorio sperimentale" di nuove tendenze, soprattutto in considerazione della mancanza di appigli normativi che guidino l'interprete nell'individuare la disciplina applicabile. (I criteri di accertamento del nesso di causalità nell'illecito omissivo: il "più probabile che non" spazza via ogni certezza, di Maria Fontana Vita della Corte, in rivista responsabilità medica)

A ben vedere, come sostenuto da autorevole dottrina, in tale ipotesi la responsabilità del medico andrà parametrata al grado di autonomia decisionale riconosciuto alla macchina. In tal senso, tanto maggiore sarà il grado di autonomia del robot intelligente e tanto più difficile sarà dimostrare un nesso causale tra il comportamento posto in essere dal medico e il danno cagionato al paziente.

Posto che, in futuro, sarà l'intelligenza artificiale stessa a suggerire ai medici i trattamenti raccomandati per il singolo paziente in base ai più recenti studi scientifici, un caso in cui certamente verrà in rilievo la responsabilità (per colpa) del medico sarà quello in cui la tecnologia intelligente abbia fornito indicazioni errate a causa dell'errato utilizzo della macchina da parte dell'operatore sanitario.¹⁷⁷

Viceversa, nell'ipotesi in cui il medico si sia affidato alla macchina nell'individuazione del trattamento più idoneo, ma la terapia non dovesse produrre i risultati sperati a causa di un errore insito nel processo decisionale e/o esecutivo della macchina, non si potrà addurre la responsabilità del professionista e, per di più, laddove si dovesse dimostrare che il medico non potesse in alcun modo accorgersi dell'errore, ci troveremo di fronte a un vero e proprio evento interruttivo del rapporto di causalità tra condotta del medico ed evento lesivo.

2.3 Cenni sulla responsabilità della struttura sanitaria

Per quanto riguarda gli ospedali, occorre innanzitutto ricordare brevemente quanto segue la responsabilità delle strutture sanitarie, come disciplinata fino ad oggi.

Secondo le linee guida stabilite dalla giurisprudenza e supportate dalla normativa in materia

¹⁷⁷ Votano G., *"Intelligenza artificiale in ambito sanitario: il problema della responsabilità civile"* cit.

leggi in materia (articolo 7, comma 1, della legge n. 24 dell'8.3.2017), il rapporto che lega le strutture sanitarie (pubbliche o private) con il paziente nasce dal contratto di "ricovero" o "assistenza sanitaria".

E' noto che le aziende sanitarie instaurano con i propri pazienti un rapporto negoziale e atipico, denominato contratto di ricovero o ospedalità, nel quale si configurano una serie di obbligazioni assunte in nome e per conto degli stessi o con la mediazione dei medici ospedalieri, che, in caso di inadempimento, diventano fonte di responsabilità ai sensi degli articoli 1218 e 1228 del codice civile.

A questo proposito, la giurisprudenza si limita a sancire che "la responsabilità delle istituzioni infermieristiche in generale riguarda la violazione dell'obbligo che l'istituzione infermieristica stessa ha di fornire direttamente al paziente l'agenzia per realizzare l'intervento richiesto".¹⁷⁸

Lo statuto (e la legge di accompagnamento 24/2017) non spiegano specificamente i criteri di attribuzione della responsabilità per violazione del dovere, ma si può essere certi che la sua natura oggettiva si applica anche alle strutture sanitarie. Anche la giurisprudenza è molto chiara al riguardo: la Corte di Cassazione ha analizzato l'attribuzione dell'onere della prova a questo proposito e, in effetti, ha espressamente affermato che il debitore di servizi deve provare "in particolare che la prestazione di servizi è divenuta impossibile (eliminazione dei fatti) a causa di motivi imprevedibili e inevitabili". La "prestazione (estinzione dei diritti)" deve essere provata.

Tuttavia, ci si chiede in che misura questi sistemi possano essere adattati ai problemi specifici posti dall'uso dei sistemi di IA nell'erogazione dell'assistenza sanitaria.

La risposta a questa domanda sembra, ad oggi, essere abbastanza semplice nel caso in cui la maggior parte della popolazione utilizza gli strumenti dell'IA solo per supportare le attività dei professionisti e assistere nel processo decisionale clinico.

Si è già avuto modo di chiarire che questo è un caso in cui non c'è un ruolo per i medici nel declassamento specifico, in quanto l'affidabilità dei risultati deve essere verificata.

e le sue possibili confutazioni, sulla base della conoscenza del sistema e delle specificità del caso specifico.

Ciò la rende un'opzione anche se il sistema produce risultati errati.

Perché gli errori non sono inevitabili, per quanto sofisticata sia la tecnologia.

Anche le applicazioni intelligenti più sofisticate sono lasciate ai professionisti, valutare la loro affidabilità ed evitare l'adesione automatica alle soluzioni fornite da una tecnologia problematica. Pertanto, le carenze sistemiche sono destinate a permanere nel prossimo futuro.

¹⁷⁸ Cass. SS.UU. 1 luglio 2002 n. 9556. Cfr. anche Cass. SS.UU. 11 gennaio 2008 n. 577

In questo caso, un'eventuale negligenza da parte del medico comporterà la responsabilità sia del medico che della casa di cura, dove la persona lavora, secondo l'orientamento unitario già citato.

Se, invece, non viene stabilito un profilo di colpa, il problema si complica.

Se non è possibile trovare un profilo di azione condannabile di un medico, ad esempio, il dipartimento dei servizi medici è lasciato alle azioni autonome delle macchine, che si guastano, causando danni al paziente.

In uno scenario del genere, è un prerequisito per la responsabilità, l'attività autonoma del sistema. La strada dell'addebito di responsabilità in capo alla Struttura sanitaria per insufficiente organizzazione, sarebbe ipotizzabile nel caso del danno derivante da cattivo funzionamento del sistema intelligente.

Come già spiegato, il rapporto complesso e insolito tra ospedali e pazienti impone anche obblighi organizzativi agli ospedali.

In caso contrario, l'ospedale può essere ritenuto responsabile ai sensi degli articoli 1218 e 1176 del Codice Civile per la violazione del dovere primario di non arrecare un danno ingiusto al paziente, non avendo predisposto le attrezzature necessarie per il corretto svolgimento della prestazione medica.

Si tratta di un caso di responsabilità che ha acquisito nuova attenzione grazie al crescente utilizzo della tecnologia, anche intelligente, a fini diagnostici, terapeutici e riabilitativi. Infatti, l'assistenza sanitaria è diventata dipendente dall'uso massiccio di apparecchiature mediche sofisticate e in particolare di dispositivi, e questo è vero nella misura in cui questi sono dotati di componenti informatici di base che in passato erano quasi inesistenti.

Ciò ha un impatto significativo sulla responsabilità delle organizzazioni sanitarie, che sono costrette a prendere in considerazione più frequentemente il rischio di "incidenti tecnici", ossia il guasto o il degrado delle caratteristiche o delle prestazioni delle apparecchiature, o la presenza di eventi avversi che portano al deterioramento della salute dei pazienti o alla morte.¹⁷⁹

Quale potrebbe essere il punto fondamentale per eliminare la responsabilità dei debitori delle strutture sanitarie?

Taluni ostacoli possono essere riscontrati nel contesto specifico dell'intelligenza artificiale, soprattutto in tutti i casi in cui la struttura sanitaria non ha il pieno controllo del dispositivo intelligente, basato sul cloud.

¹⁷⁹ Cerea Francesca, *“I danni da Intelligenza Artificiale: una nuova frontiera”* Università degli studi di Bergamo, 2020/2021.

Si pensi all'ipotesi dell'intelligenza artificiale basata sul cloud sviluppata per le attività diagnostiche che è controllata dal suo stesso produttore e che non riesce a diagnosticare correttamente, ritardando così il trattamento e causando danni ingiustificati. Se l'algoritmo fallisce a causa di una programmazione errata, di aggiornamenti della base di conoscenze non eseguiti dal produttore, di uno sviluppo autonomo anomalo che non può essere impedito da uno spegnimento di emergenza o per altri motivi simili, il fallimento può essere considerato causato da fattori al di fuori del controllo della struttura sanitaria debitrice.

Più in generale, si applica un comune paradigma di responsabilità contrattuale. Se per essere assolto, il debitore deve provare che gli eventi che hanno impedito l'esecuzione accurata del servizio previsto non erano a lui imputabili nel minimo, ma erano oggettivi e al di fuori del suo controllo, allora la struttura sanitaria che utilizza l'intelligenza artificiale deve provare che questi eventi non sono a lui imputabili nel minimo, deve provare che erano oggettivi e al di fuori del suo controllo.

Che fossero oggettivi e al di fuori del suo controllo significa che la struttura sanitaria che utilizza l'intelligenza artificiale deve dimostrare che l'errore algoritmico era inevitabilmente al di fuori del suo controllo: pertanto, non deve essere in grado di influenzare i dati a disposizione dell'IA, i suoi metodi di elaborazione e i suoi aggiornamenti. Si deve inoltre presumere che tali errori siano del tutto inaspettati:

Nel tempo, sarà difficile per una struttura sanitaria considerare "non provocata" una violazione che danneggia un paziente causata da un dispositivo medico di IA che ha avvisato il paziente della necessità di manutenzione o di aggiornamenti, o che ha avvisato il paziente della necessità di manutenzione o di aggiornamenti.

Tuttavia, non vi è alcuna violazione del contratto se un utente non professionale che ha ricevuto un'apparecchiatura medica intelligente da una struttura sanitaria (ad esempio, la cosiddetta assistenza domiciliare) e ha ricevuto istruzioni appropriate dalla struttura sanitaria riduce l'efficacia dei suoi servizi.

Questi principi sono appropriati anche per le ricostruzioni che considerano l'apparecchiatura medica autonoma come un "aiuto" alla struttura medica: esiste sempre un potenziale legame tra l'attività della macchina intelligente e un malfunzionamento, che può portare a usi imprevisti, completamente inappropriati e non previsti dalla struttura obbligata.

Inoltre, questi principi sono appropriati anche per la ristrutturazione dei dispositivi medici autonomi come "dispositivi di assistenza" per le strutture sanitarie. Poiché esiste sempre un necessario nesso di opportunità tra l'attività di un dispositivo intelligente e il danno che esso arreca all'ambito medico (uso impossibile al di fuori delle attività proprie della struttura), la struttura

dovrà tipicamente soddisfare le disposizioni combinate delle sezioni 1228 e 1218 del Codice. Una situazione davvero difficile da raggiungere è il nesso (rilevanza).

Il nesso può essere interrotto se l'IA viene utilizzata senza autorizzazione in condizioni o modalità non approvate dalla struttura medica che la utilizza. Il nesso si considera interrotto se si ritiene che vi sia l'opportunità necessaria per un uso lecito e consensuale del dispositivo algoritmico fornito dalla struttura sanitaria.¹⁸⁰

Ovviamente, se il medico non è responsabile, non è esclusa allo stesso modo la responsabilità della struttura sanitaria. Finché quest'ultima ha il controllo dell'IA, la struttura sanitaria può essere responsabile, secondo le norme in corso di rapida analisi, anche se non c'è negligenza da parte del medico.

Inoltre, si è già detto che la "struttura duale" della responsabilità stabilita dalla legge Gelli-Bianco è perfettamente compatibile con questa ipotesi (anzi, idealmente la favorisce, e il regime di responsabilità per i medici è più attenuato di quello ipotizzato per le strutture sanitarie e quindi della responsabilità del medico).¹⁸¹

Inoltre, le strutture sanitarie sono, per volontà del legislatore, destinatarie selettive di cause per il risarcimento dei pazienti danneggiati). In generale, la posizione delle strutture sanitarie rispetto allo standard di imputazione della responsabilità per violazione del contratto è destinata a rimanere invariata.

Anche se il medico viene ritenuto il maggior responsabile, la struttura può essere ritenuta responsabile per la responsabilità della struttura stessa e per i danni da essa causati. D'altro canto, se l'errore algoritmico è causato da fattori completamente estranei alla sfera di influenza della struttura, quest'ultima può evitare la responsabilità (senza ridurre la probabilità di responsabilità del produttore) ricorrendo allo scudo della non attribuzione.

¹⁸⁰ Esposito Leonardo, “*Strumenti autonomi e responsabilità civile con particolare riguardo al settore medico*” Università LUISS Dipartimento di Giurisprudenza, 2020/2021;

¹⁸¹ Cerea Francesca, “*I danni da Intelligenza Artificiale: una nuova frontiera*” Università degli studi di Bergamo, 2020/2021.

3. La recentissima proposta del Parlamento europeo e del Consiglio relativa all'adeguamento delle norme in materia di responsabilità civile extracontrattuale all'intelligenza artificiale (direttiva sulla responsabilità da intelligenza artificiale)

Dall'analisi fino ad ora condotta è emerso che nel caso di eventi avversi nella prestazione di cure attraverso l'uso di tecnologie emergenti, l'IA e la robotica avanzata sembrano essere riconosciute a livello nazionale come elementi neutri che, di per sé, dovrebbero essere inclusi nell'ambito dell'obbligazione di mezzi disciplinata dal principio della responsabilità per negligenza in relazione alla natura delle attività caratteristiche degli operatori sanitari. L'uso di tecnologie emergenti può aumentare la precisione e l'acutezza degli interventi medici, ma questo non porta a caratterizzare i servizi medici come un obbligo di risultato.

Pertanto, l'insuccesso del trattamento o il peggioramento delle condizioni di salute di un paziente a seguito dell'utilizzo di A.I. non implica automaticamente l'esistenza di una negligenza medica. In sintesi, gli operatori sanitari restano in regime di responsabilità per negligenza nella prestazione di servizi sanitari, secondo quanto previsto dalla Legge Gelli-Bianco, anche nel caso di utilizzo di nuove tecnologie o di qualsiasi strumento, applicazione o dispositivo medico.¹⁸²

Come è chiaro è ancora lunga la strada per una definizione pacifica e condivisa per un regime di responsabilità per danni cagionati dall'utilizzo dell'I.A.

Tuttavia un enorme passo in avanti è stato effettuato di recente con la proposta del Parlamento europeo e del Consiglio relativa all'adeguamento delle norme in materia di responsabilità civile extracontrattuale all'intelligenza artificiale.

In data 28 settembre 2022, la Commissione europea ha presentato una proposta legislativa per una nuova direttiva volta ad adeguare le norme in materia di responsabilità civile extracontrattuale all'intelligenza artificiale.¹⁸³

¹⁸² Colaruotolo, A. *“Intelligenza artificiale e responsabilità medica: novità, continuità e criticità”* in *Responsabilità medica, diritto e pratica clinica*, 2022, v. 3, p. 299.

¹⁸³ Si veda la Direttiva del parlamento europeo e del consiglio relativa all'adeguamento delle norme in materia di responsabilità civile extracontrattuale all'intelligenza artificiale (direttiva sulla responsabilità da intelligenza artificiale), 2022, <https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:52022PC0496&from=EN>

Le nuove norme renderanno più facile, il risarcimento per coloro che sono stati discriminati in procedure di assunzione che coinvolgono tecnologie basate sull'IA.

Ad esempio , con riguardo alla responsabilità del prodotto, proteggeranno i consumatori e incoraggeranno l'innovazione.

In primo luogo, propone di modernizzare le norme esistenti sulla responsabilità oggettiva del produttore per i prodotti difettosi (dalle tecnologie intelligenti ai farmaci).

Le norme riviste garantiranno la certezza del diritto per le imprese, consentiranno loro di investire in prodotti nuovi e innovativi e assicureranno che le vittime ricevano un equo risarcimento per i danni causati da prodotti difettosi, compresi quelli digitali e rigenerati.

In secondo luogo, per la prima volta, la Commissione propone di armonizzare le norme nazionali in materia di responsabilità per l'intelligenza artificiale (IA), rendendo più semplice il risarcimento dei danni causati dall'IA.

L'obiettivo della direttiva sulla responsabilità per l'IA è migliorare il funzionamento del mercato interno stabilendo requisiti uniformi per la responsabilità civile extracontrattuale per i danni causati con il coinvolgimento di sistemi di IA.

L'obiettivo generale della proposta è quello di promuovere l'introduzione di un'IA affidabile e di sfruttarne appieno i vantaggi per il mercato interno, garantendo che le vittime di danni causati dall'IA ottengano una protezione equivalente a quella delle vittime di danni causati da prodotti in generale.¹⁸⁴

La proposta mira inoltre a ridurre l'incertezza giuridica per le imprese che sviluppano o utilizzano l'IA per quanto riguarda la loro possibile esposizione alla responsabilità e a prevenire l'emergere di adattamenti frammentari specifici per l'IA delle norme nazionali sulla responsabilità civile.

La base giuridica della proposta è l'articolo 114 del TFUE, che prevede l'adozione di misure volte a garantire l'istituzione e il funzionamento del mercato interno.

La scelta di una direttiva lascia agli Stati membri una certa flessibilità nel recepimento interno della legislazione, poiché le norme direttamente applicabili sarebbero troppo rigide in relazione all'ambito della responsabilità civile, che si basa su tradizioni giuridiche specifiche e consolidate in ogni Stato membro.

La nuova proposta di direttiva sulla responsabilità per l'IA mira ad armonizzare le norme sulla responsabilità civile extracontrattuale per i danni causati dai sistemi di intelligenza artificiale (articolo 1).

¹⁸⁴ Cit.

La suindicata direttiva non definirebbe l'IA, ma farebbe riferimento allo stesso concetto generale di IA contenuto nell'AI Act e in particolare alla sua definizione di "sistemi di IA".¹⁸⁵

Le nuove norme si applicherebbero ai danni causati dai sistemi di IA, indipendentemente dal fatto che siano definiti o meno ad alto rischio ai sensi dell'AI Act.

La direttiva sulla responsabilità per l'IA riguarda le norme sulla responsabilità civile "extracontrattuale", ossia le norme che prevedono una richiesta di risarcimento a prescindere da un legame contrattuale tra la vittima e il responsabile.¹⁸⁶

Le norme garantirebbero che qualsiasi tipo di vittima (individui o imprese) possa essere risarcita se viene danneggiata da una colpa o da un'omissione di un fornitore, di uno sviluppatore o di un utente dell'IA che provoca un danno coperto dal diritto nazionale (ad esempio, salute, proprietà, privacy, ecc.) o per i danni cagionati, in ambito sanitario, alla salute del paziente dall'uso di sistemi di supporto alla diagnosi medica che, a causa della scarsa rappresentatività o del mancato aggiornamento dei dati che gli siano stati forniti, segnalino, come urgente, un erroneo trattamento curativo o farmacologico ovvero non indichino un trattamento sanitario necessario.¹⁸⁷

La direttiva UE sulla responsabilità per l'IA segue il lavoro definito nella legge UE sull'IA per garantire la coerenza. Più precisamente, l'art. 2 della Direttiva prevede che le azioni di risarcimento danni possono essere avanzate, oltre che dal danneggiato, anche da soggetti succeduti o surrogati nei diritti del danneggiato, quali una compagnia di assicurazioni o gli eredi di una vittima deceduta. Inoltre, l'articolo 2, comma 6, lettera c), prevede che l'azione risarcitoria può essere promossa anche da chi agisce per conto di una o più persone lese, in conformità al diritto comunitario o nazionale di uno Stato membro dell'Unione.

La Direttiva introduce anche un meccanismo specifico per la divulgazione delle prove al fine di "fornire alle persone che richiedono un risarcimento per danni causati da sistemi di IA ad alto rischio mezzi efficaci per identificare le persone potenzialmente responsabili e prove pertinenti per

¹⁸⁵ I sistemi di IA sono definiti nella bozza dell'atto di IA come software sviluppato con determinate tecniche e approcci (apprendimento automatico e approcci basati sulla conoscenza, approcci statistici, ecc.) in grado, per un dato insieme di obiettivi definiti dall'uomo, di generare output come contenuti, previsioni, raccomandazioni o decisioni che influenzano gli ambienti con cui questi sistemi interagiscono.

¹⁸⁶ Per approfondimento di veda: Domande e risposte: direttiva sulla responsabilità da intelligenza artificiale, Commissione Europea, 2022, https://ec.europa.eu/commission/presscorner/detail/en/QANDA_22_5793;

¹⁸⁷ Burylo, Y. (2022). CIVIL LIABILITY FOR DAMAGE CAUSED BY ARTIFICIAL INTELLIGENCE: THE MODERN EUROPEAN APPROACH. *Entrepreneurship, Economy and Law*, (6), 5–11. <https://doi.org/10.32849/2663-5313/2022.6.01>

una richiesta di risarcimento”.¹⁸⁸ Questo mira ad aiutare le vittime ad accedere alle prove pertinenti. Infatti, l'articolo 3, paragrafo 1, della direttiva lo prevede "*Gli Stati membri provvedono affinché i tribunali nazionali siano autorizzati, su richiesta di un potenziale attore che ha precedentemente chiesto a un prestatore, a una persona soggetta agli obblighi di un prestatore ai sensi [dell'articolo 24 o dell'articolo 28, paragrafo 1, della legge sull'IA] o un utente a divulgare le prove pertinenti a sua disposizione su uno specifico sistema di IA ad alto rischio sospettato di aver causato danni, ma a cui è stato rifiutato, o un ricorrente, di ordinare la divulgazione di tali prove a tali persone*".

Inoltre, la direttiva UE sulla responsabilità per l'IA introduce una "presunzione di causalità" confutabile. Secondo tale principio, il giudice deve presumere che l'inosservanza di un determinato obbligo rilevante per il danno causato dal sistema di IA, se il nesso di causalità è probabile, abbia causato il danno.

La direttiva sulla responsabilità dell'IA non influirebbe sulle norme esistenti stabilite in altri atti legislativi dell'UE, in particolare sulle norme dell'UE che regolano le condizioni di responsabilità nel settore dei trasporti, sulla proposta di revisione della direttiva sulla responsabilità dei prodotti o sulla legge sui servizi digitali.

Inoltre, sebbene la direttiva sulla responsabilità per l'IA non si applichi alla responsabilità penale, potrebbe essere applicabile alla responsabilità statale, dato che le autorità statali sono soggette agli obblighi previsti dalla legge sull'IA.

La proposta di revisione della direttiva mira a modernizzare la esistente nell'UE, basato sull'assenza di colpa (responsabilità oggettiva) per i prodotti e si applicherebbe alle richieste di risarcimento avanzate da privati contro il produttore per i danni causati da prodotti difettosi.¹⁸⁹

La nuova direttiva sulla responsabilità civile per danno da prodotti difettosi propone invece una riforma mirata dei regimi nazionali di responsabilità per colpa nazionali di responsabilità per colpa e si applicherebbe alle richieste di risarcimento, avanzate da qualsiasi persona fisica o giuridica nei confronti di qualsiasi persona, per difetti che influenzano il sistema di IA che ha causato il danno. La Commissione europea intende che la direttiva sulla responsabilità dell'IA e la direttiva sulla responsabilità del prodotto riveduta si integrino a vicenda nella costruzione di un quadro di

¹⁸⁸ Direttiva del parlamento europeo e del consiglio relativa all'adeguamento delle norme in materia di responsabilità civile extracontrattuale all'intelligenza artificiale (direttiva sulla responsabilità da intelligenza artificiale), 2022, <https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:52022PC0496&from=EN>

¹⁸⁹ Lusardi Giacomo, "Danni causati dall'intelligenza artificiale, chi paga? Cosa prevede la proposta di direttiva Europea" Agendadigitale.eu, 2022, <https://www.agendadigitale.eu/cultura-digitale/danni-causati-dallintelligenza-artificiale-chi-paga-cosa-prevede-la-proposta-di-direttiva-ue/>

responsabilità per i sistemi di IA e i prodotti abilitati all'IA, insieme al più ampio quadro normativo ai sensi della legge sull'IA. Questa serie di normative dell'UE in materia di IA si trova attualmente in varie fasi del processo legislativo dell'UE e sarà soggetta a ulteriori negoziati e potenziali modifiche prima di entrare in vigore.¹⁹⁰

Rimangono interrogativi in termini di come la responsabilità oggettiva e i regimi basati sulla colpa opereranno fianco a fianco nella pratica; in che modo questi regimi di responsabilità si collegheranno nella pratica al quadro di conformità proposto dalla legge sull'IA per i sistemi di IA; e come si svilupperanno le pratiche di mercato per l'attribuzione della responsabilità lungo l'IA e le catene di fornitura dei dati.

È probabile che ci vorranno anni prima che la proposta diventi legge, dopo aver potenzialmente subito importanti modifiche al suo testo, poiché deve ancora essere sottoposta all'iter legislativo dell'UE. Tuttavia, insieme alla bozza dell'EU AI Act (che dovrebbe diventare legge nel 2023 o 2024), la direttiva dimostra un chiaro impulso nel mondo legale per l'adozione di norme severe per l'IA, nell'UE e all'estero.

Conclusioni

Dall'analisi fin qui realizzata è possibile affermare che l'attuale legislazione è insufficiente a garantire la tutela delle vittime dell'uso di dispositivi di IA nel settore sanitario sotto due aspetti: in particolare, l'identificazione degli attori coinvolti nella catena di produzione (o nell'implementazione dei servizi) e la riduzione dell'onere della prova da parte delle vittime, poiché la rivoluzione dell'IA comporta una situazione di opacità e pluralismo dei sistemi.

Quanto maggiore è il numero di attori coinvolti nel ciclo, tanto più difficile diventa il suo svolgimento.

È impossibile dimostrare che un particolare incidente sia causalmente legato a una particolare decisione umana.

¹⁹⁰ Marino David Maria e Olivieri Andrea, “*Chi paga per i danni causati dall'intelligenza artificiale?*”

In assenza di una regolamentazione specifica sull'argomento, ecco quali sono le norme e i casi recenti a cui si fa riferimento in questo ambito” Wired, 2023, <https://www.wired.it/article/intelligenza-artificiale-responsabilita-regole/>.

Le decisioni umane vengono prese durante la progettazione, la produzione o il funzionamento dei dispositivi intelligenti.

Questo onere diventa ancora più gravoso se si considera che tutti i sistemi di IA sono collegati ad altri sistemi (basati o meno sull'IA), si basano su dati esterni e sono esposti al rischio di violazioni della sicurezza informatica. Di conseguenza, potrebbero verificarsi diversi casi in cui i pazienti danneggiati dai dispositivi di IA non possono dimostrare la negligenza del produttore o del distributore.

degli utenti del dispositivo stesso, oppure quando non è possibile correlare il danno ad un difetto nel prodotto.

Questo ci porta alla conclusione che i regimi di responsabilità basati sulla negligenza sono considerati appropriati nel contesto della IA quando la natura autonoma del dispositivo esclude il coinvolgimento diretto dell'uomo nella causazione del danno¹⁹¹.

Per questo motivo, la teoria accademica prevalente considera la responsabilità oggettiva come il sistema più appropriato per regolare tali situazioni¹⁹². In base a questo sistema, una persona che crea, mantiene, controlla o interferisce con un sistema di IA può essere ritenuta responsabile per i danni causati dal dispositivo, sulla base del fatto che una persona che introduce un rischio per il pubblico è responsabile se il danno deriva da tale rischio.

A sostegno di queste conclusioni, è stato sottolineato che i regimi di responsabilità oggettiva riducono i costi associati all'onere della prova in controversie complesse¹⁹³ e che i regimi di responsabilità oggettiva possono incoraggiare l'innovazione e migliorare la sicurezza dei prodotti se forniscono un quadro giuridico sicuro e prevedibile all'interno del quale le aziende possono operare.

Un'altra scuola di pensiero sostiene che l'applicazione di regimi di responsabilità oggettiva costituisce un'importante barriera all'ingresso di nuove imprese nel mercato perché, a differenza delle grandi imprese con risorse significative, le piccole imprese non possono sostenere l'onere finanziario di tali regimi¹⁹⁴.

¹⁹¹Rachum-Twaig, "Whose robots is it anyway?: liability of artificial intelligence-based robots", in *University of Illinois Law Review*, 2020, 1141

¹⁹² V. A. Lior, "AI Entities as AI Agents: Artificial Intelligence Liability and the AI Respondeat Superior Analogy", in *Mitchel Hamline Law Review*, vol. 46, 2020, 1043;

¹⁹³ D.C. Vladeck, "Machines Without Principals: Liability Rules and Artificial Intelligence", in *Washington Law Review*, vol. 89, 117, 2014

¹⁹⁴ Y. Bathae, "The Artificial Intelligence Black Box and the Failure of Intent and Causation", in *Harvard Journal of*

Tuttavia, il dibattito accademico prevalente respinge questa tesi, sottolineando che il diritto assicurativo svolge un ruolo importante in questo contesto come strumento di copertura del rischio volto a incoraggiare le imprese ad assumere rischi in nuove aree tecnologiche.

Le difficoltà legate alla determinazione del responsabile in caso di danno non possono quindi essere risolte dal diritto assicurativo.

Le difficoltà legate alla determinazione del responsabile in caso di danno non possono essere risolte dal diritto assicurativo. Questo problema può essere superato istituendo un sistema di responsabilità oggettiva e di assicurazione obbligatoria (insieme alla creazione dell'auspicato fondo di garanzia) che coinvolga tutti gli attori della catena di produzione, commercializzazione e utilizzo dei dispositivi intelligenti.

Nel prossimo futuro, le compagnie assicurative saranno comunque chiamate a precisare se le relative polizze stipulate con strutture sanitarie e i professionisti coprono o meno i danni derivanti da utilizzo di sistemi intelligenti.

Presumibilmente, l'utilizzo di un modello di risarcimento "senza colpa" ha il vantaggio di ridurre il contenzioso nei casi di controversie legali che coinvolgono interessi sensibili e un'insufficiente internalizzazione dei costi. Si tratta di un meccanismo sperimentato nell'ordinamento giuridico italiano in relazione ad attività ad alto impatto sociale, come gli incidenti stradali senza colpa e le politiche di sanità pubblica.

All'esito dell'indagine svolta, dal punto di vista *de jure condendo*, la soluzione più appropriata, senza snaturare l'ordinamento giuridico, sarebbe quella di assimilare le raccomandazioni di IA formulate dai sistemi di supporto alle decisioni cliniche in linee guida, con sanzioni per la responsabilità in caso di deviazioni ingiustificate da parte degli operatori sanitari. Questa soluzione renderebbe gli operatori sanitari consapevoli e responsabili dell'uso di IA nella pratica medica, senza modificare lo status della legge Gelli-Bianco, che limita esplicitamente la loro responsabilità all'articolo 2043 del Codice Civile ed esclude altre forme di responsabilità.

Questo approccio avrebbe anche il vantaggio di proteggere l'autonomia degli operatori sanitari nelle decisioni professionali, evitando tecniche mediche difensive e preservando la specificità dei casi particolari.

BIBLIOGRAFIA

- “*Artificial Intelligence Improves Coronary Artery Disease Detection, New Study Finds.*” Home, Retrieved, 2023, si veda <https://www.bsecho.org/Public/Public/News/Articles/2021/2021-12/202112-Ultromics.aspx>
- Ascolese Raffaella, “*La natura giuridica della responsabilità medica: il panorama legislativo dopo la Legge n. 24 del 2017*” D.it, 21/03/2022, <https://www.diritto.it/la-natura-giudica-della-responsabilita-medica-il-panorama-legislativo-dopo-la-legge-n-24-del-2017/>
- Azzali, E. “*Notes on Recent European Commission Plans for regulation of Artificial Intelligence.*” In *Management* 20, 2020 pp. 20- 22.
- Belk, Russell. “*Ethical issues in service robotics and artificial intelligence.*” *The Service Industries Journal*, 2021 pp. 860-876.
- Bezaire J, Felton KW, Greve P, Allen D. “*Hospital negligent credentialing liability*” 2017 consultabile al <https://www.willistowerswatson.com/en-US/Insights/2017/08/hospital-negligentcredentialing-liability>
- Blasimme, Alessandro “*Governare la biomedicina: l'etica della ricerca scientifica come questione pubblica*” Paradigmi, 2011
- Cerea Francesca, “*I danni da Intelligenza Artificiale: una nuova frontiera*” Università degli studi di Bergamo, 2020/2021;
della responsabilità civile
- Chinè G., Fratini M., Zoppini A., “*Manuale di diritto civile*”, Molfetta (BA), Nel diritto editore, 2017, IX edizione p. 2181;
- Ciancimino M., “*AI-Based Decision-Making Process in Healthcare*”, (2022), 11, *Journal of European Consumer and Market Law*, Issue 5, pp. 173-180, <https://kluwerlawonline.com/JournalArticle/Journal+of+European+Consumer+and+Market+Law/11.5/EuCML2022029>
- Cohen IG, Mello MM. “*HIPAA and protecting health information in the 21st century*” *JAMA* 2018 <https://doi.org/10.1001/jama.2018.5630>
- Cohen, I. Glenn, and Michelle M. Mello. “*Big data, big tech, and protecting patient privacy.*” *Jama* 2019: pp. 1141-1142.
- Colaruotolo, A. “*Intelligenza artificiale e responsabilità medica: novità, continuità e criticità*” in *Responsabilità medica, diritto e pratica clinica*, 2022, v. 3, p. 299;

- Coletti Elisa, “*Intelligenza artificiale e attività sanitaria. profili giuridici dell’utilizzo della robotica in medicina*” in Rivista di Diritto dell’Economia, dei Trasporti e dell’Ambiente, 2021, Vol. XIX, pp. 201-219;
- Comunicazione della commissione al Parlamento europeo, al Consiglio europeo, al Consiglio, al Comitato economico e sociale europeo e al comitato delle regioni - piano coordinato sull'intelligenza artificiale” consultabile al <https://eur-lex.europa.eu/legal-content/IT/TXT/?uri=CELEX%3A52018DC0795>
- Corti, Quando L'intelligenza Artificiale Salva Vite Umane.” 01net, 2019, si veda <https://www.01net.it/corti-intelligenza-artificiale-salva-vite-umane/>
- De Paz, Juan F., et al. "An integrated system for helping disabled and dependent people: AGALZ, AZTECA, and MOVI-MAS projects." Springer, Cham, 2015
- De Vinco A., “L’ utilizzo degli M.M.S. e la protezione dei dati personali nel Regno Unito” AmbienteDiritto, 2006, si veda www.ambientediritto.it/dottrina/Dottrina_2006/utilizzo_M.M.S_devinco.htm
- Der Tagesspiegel “EU guidelines: Ethics Washing Made in Europe” Aktuelle News: Nachrichten aus Berlin und der Welt. 2019, consultabile al <https://www.tagesspiegel.de/politik/ethics-washing-made-in-europe-5937028.html>
dium.com/@virginiadignum/the-art-of-ai-accountability-responsibility-transparency-48666ec92ea5.
- Ebers, M., Hoch, V. R., Rosenkranz, F., Ruschemeier, H., & Steinrötter, B. “The European Commission’s proposal for an artificial intelligence act—a critical assessment by members of the Robotics and AI Law Society” 2021, 589–603 si veda <https://doi.org/10.3390/j4040043>.
- Esposito Leonardo, “Strumenti autonomi e responsabilità civile con particolare riguardo al settore medico” Università LUISS Dipartimento di Giurisprudenza, 2020/2021;
- Ethics Guidelines for Trustworthy Ai.” Shaping Europe's Digital Future, consultabile <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>.
- Executive Office of the President. Memorandum for the Heads of Executive Departments and Agencies, 2018, M-18-22, <https://www.whitehouse.gov/wp-content/uploads/2018/07/M-18-22.pdf>
- Fasan, Marta “Oms – Ethics Ad Governance of Artificial Intelligence for Health: Who Guidance.” Biodiritto, 2021, <https://www.biodiritto.org/Biolaw-pedia/Docs/OMS-Ethics-ad-Governance-of-Artificial-Intelligence-for-Health-WHO-Guidance>

- Fei, H., et al. “*A Planning and Scheduling Problem for an Operating Theatre Using an Open Scheduling Strategy*” *Computers & Industrial Engineering*, 2010 vol. 58, n. 2, pp. 221–230., <https://doi.org/10.1016/j.cie.2009.02.012>.
- Fioretti Marta, “*Questioni giuridico-dogmatiche sull’intelligenza artificiale*” *SalvisJuribus*, 16/08/2021, <http://www.salvisjuribus.it/questioni-giuridico-dogmatiche-sullintelligenza-artificiale/>
- Floridi, Luciano. “*Establishing the Rules for Building Trustworthy AI*” in *Electronic Journal*, 2019, pp. 261–262, <https://doi.org/10.1038/s42256-019-0055-y>
- Fusaro Arianna, “*Quale modello di responsabilità per la robotica avanzata? Riflessioni a margine del percorso europeo*” *La nuova giurisprudenza civile commentata*, 2020;
- Gadmin. “*Intelligenza Artificiale in Medicina, L’esempio Di Deepmind in Oculistica.*” Galeno, 2020, <https://cassagaleno.eu/intelligenza-artificiale-in-medicina-lesempio-di-deepmind-in-oculistica/>
- Gailhofer, P.; Herold, A.; Schemmel, Giappone; Scherf, CS-S.; Urruzia, C.; Kohler, A.; Braungardt, S. *The Role of Artificial Intelligence in the European Green Deal*, studio richiesto dalla commissione AIDA del Parlamento europeo, 2021 consultabile al [https://www.europarl.europa.eu/RegData/etudes/STUD/2021/662906/IPOL_STU\(2021\)662906_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2021/662906/IPOL_STU(2021)662906_EN.pdf)
- Gerke, Sara, et al. “*Ethical and Legal Challenges of Artificial Intelligence-Driven Healthcare.*” *Artificial Intelligence in Healthcare*, 2020, pp. 295–336.
- Gervais, D. “*Is Intellectual Property Law Ready for Artificial Intelligence?*” *GRUR International*, 2020 vol. 69, n. 2, pag. 117–118, <https://doi.org/10.1093/grurint/ikz025>;
- Gryz, Jarek, Marcin Rojszczak. “*Black box algorithms and the rights of individuals: No easy solution to the "explainability" problem.*” *Internet Policy Review*, 2021 vol. 10, n. 2, 2021, pag. 1-24 <https://doi.org/10.14763/2021.2.1564>;
- In assenza di una regolamentazione specifica sull’argomento, ecco quali sono le norme e i casi recenti a cui si fa riferimento in questo ambito” *Wired*, 2023, <https://www.wired.it/article/intelligenza-artificiale-responsabilita-regole/>
- Iqbal, Sajid, et al. “*Application of Intelligent Agents in Health-Care: Review.*” *Artificial Intelligence Review*, 2016 vol. 46, n. 1, pp. 83–112., <https://doi.org/10.1007/s10462-016-9457-y>.
- J. Becker “*Insufficient Protections for Health Data Privacy: Lessons from Dinerstein v. Google*”, <https://blog.petrieflom.law.harvard.edu/>
- Journal, Hipaa. “*De-identification of Protected Health Information: 2023 Update.*” *HIPAA Journal*, 2023, www.hipaajournal.com/de-identification-protected-health-information

- Karthikeyan J., Ng Yu Jin, E Ting Su Hie “*Learning Outcomes of Classroom Research, New Delhi*” L’Ordine Nuovo Publication, 2022;
- Lusardi Giacomo, “*Danni causati dall’intelligenza artificiale, chi paga? Cosa prevede la proposta di direttiva Europea*” Agendadigitale.eu, 2022, <https://www.agendadigitale.eu/cultura-digitale/danni-causati-dallintelligenza-artificiale-chi-paga-cosa-prevede-la-proposta-di-direttiva-ue/>;
- la relazione della commissione al parlamento europeo, al consiglio e al comitato economico e sociale europeo, Relazione sulle implicazioni dell'intelligenza artificiale, dell'Internet delle cose e della robotica in materia di sicurezza e di responsabilità, <https://op.europa.eu/it/publication-detail/-/publication/4ce205b8-53d2-11ea-aece-01aa75ed71a1/language-it/format-PDF> ;
- Lagioia, F. “*L’intelligenza Artificiale in Sanità: Un’analisi Giuridica*”, Giappichelli, 2020
- Lagioia, F., et al. “Automated detection of unfair clauses in online consumer contracts.” Legal Knowledge and Information Systems: JURIX 2017: The Thirtieth Annual Conference, 2017, Vol. 302 p. 145.
- Leibowitz, C. “Arterys: Deep Learning for Medical Imaging. Demystifying” Big Data and Machine Learning for Healthcare, 2017 pp. 169–173.
- Lusa. “*Hospital Do Barreiro Contesta Judicialmente Coima De 400 Mil Euros De Comissão De Dados.*” Público, 2018, si veda <https://www.publico.pt/2018/10/22/sociedade/noticia/hospital-barreiro-contesta-judicialmente-coima-400-mil-euros-comissao-dados-1848479>
- Maggino, Filomena, and Gabriele Cicerchia. “*Algoritmi, etica e diritto*” in Diritto Dell’Informazione e Dell’Informatica (II), 2019, vol. 3, p. 1161.
- Mauloni S., Mazzanti M., Buscemi L. “*Focus: nuove tecnologie e risvolti medico legali - la nuova era dell’intelligenza artificiale: profili medico-legali - the new era of artificial intelligence: medico-legal aspects*” Rivista Italiana di Medicina Legale (e del Diritto in campo sanitario), fasc.3, 1 giugno 2022, pag. 679;
- Mauloni S., Mazzanti M., Buscemi L., “*Nuove tecnologie e risvolti medico legali – la nuova era dell’intelligenza artificiale: profili medicolegali - the new era of artificial intelligence: medicolegal aspects*” in Rivista Italiana di Medicina Legale (e del Diritto in campo sanitario), 2022, fasc. 3 pag. 679;
- Marino David Maria e Olivieri Andrea, “Chi paga per i danni causati dall’intelligenza artificiale?”
- McCarthy J. “*A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*” in AI Magazine, 2006, vol. 27, ser. 4;

- Mittelstadt, Brent Daniel, and Luciano Floridi. “*The Ethics of Big Data: Current and Foreseeable Issues in Biomedical Contexts*” in Law, Governance and Technology Series, 2016, pp. 445–480, https://doi.org/10.1007/978-3-319-33525-4_19
- Moreno A. Nealon J. L. “*Applications of Software Agent Technology in the Health Care Domain*” Birkhauser Verlag, 2003.
- Matricardi Antonella, *La responsabilità Medica*, Altalex, 10/05/2018, disponibile al <https://www.altalex.com/documents/altalexpedia/2018/05/07/responsabilita-medica> ;
- N. Musacchio, G. Guaita, A. Ozzello, M.A. Pellegrini, P. Ponzani, R. Zilich, A. De Micheli “Intelligenza Artificiale e Big Data in ambito medico: prospettive, opportunità, criticità.” in The Journal of AMD, 2018, Vol. 21-3, pag. 204;
- Nwana, Hyacinth S. “Software Agents: An Overview.” The Knowledge Engineering Review, 1996 vol. 11, n. 3, pp. 205–244., <https://doi.org/10.1017/s026988890000789x>.
- National Archives and Records Administration, <https://trumpwhitehouse.archives.gov/presidential-actions/executive-order-maintaining-american-leadership-artificial-intelligence/>.
- Nesrine Kraiem, “*Intelligenza artificiale in medicina: quale impatto sulla responsabilità medica?*” 4C Legal, 16/11/2020, <https://www.4clegal.com/vivi-lacademy/intelligenza-artificiale-medicina-impatto-responsabilita-medica>;
- Noura Benhajji. “*Système multi-agents de pilotage réactif des parcours patients au sein des systèmes hospitaliers. Gestion et management*”, Université de Lorraine, 2017
- Pasceri, G. “*Intelligenza Artificiale, Algoritmo E Machine Learning: La Responsabilità Del Medico e Dell'amministrazione Sanitaria.*” Giuffrè Francis Lefebvre, 2021;
- Pattelli, Ivano. “*Privacy, Sicurezza Ed Etica Nell'utilizzo Dei Big Data e Dell'intelligenza Artificiale.*” ICT Security Magazine, 2022, si veda <https://www.ictsecuritymagazine.com/articoli/privacy-sicurezza-ed-etica-nellutilizzo-dei-big-data-e-dellintelligenza-artificiale/>
- Pier Giuseppe Monateri e Davide Gianti, “*Nesso di causalità (dir. civile)*” Diritto online, 2016, [https://www.treccani.it/enciclopedia/nesso-causale-dir-civ_\(Diritto-on-line\)/](https://www.treccani.it/enciclopedia/nesso-causale-dir-civ_(Diritto-on-line)/);
- Povyakalo, Andrey A., et al. "How to Discriminate between Computer-Aided and Computer-Hindered Decisions" in Medical Decision Making, 2013, vol. 33, n. 1, pp. 98-107.
- Price II WN, Gerke S, Cohen IG. “*Potential liability for physicians using artificial intelligence*” JAMA 2019; pp. 322
- Quintarelli, S., et al. “*An Ethical Perspective on Artificial Intelligence: Principles, Rights and Recommendations/Una Prospettiva Etica Sull'intelligenza Artificiale: Principi, Diritti E*

Raccomandazioni” in BIOLAW JOURNAL, 2022, consultabile al <https://hdl.handle.net/11311/1172731>.

- Rahimi, A. (2017). NIPS 2017 Test-of-Time Award presentation. Conference on Neural Information Processing Systems. [https://www.youtube.com/watch?v=ORHFOnaEzPc](https://www.youtube.com/watch?v=ORHFOnaEzPc;);
- Rahmat, Mohammad Hafidz, et al. “Agent-Based Modelling and Simulation of Emergency Department Re-Triage.” Business Engineering and Industrial Applications Colloquium (BEIAC), 2013, <https://doi.org/10.1109/beiac.2013.6560119>.
- Ratti, M. “La Convenzione sull'uso delle Comunicazioni Elettroniche: Le principali disposizioni.”, in Diritto dell'informatica, 2017, p. 71, si veda <https://hdl.handle.net/11585/580169>
- Regolamento Generale Sulla Protezione Dei Dati. Home - Garante Privacy, si veda www.garanteprivacy.it/documents/10160/0/Regolamento+UE+2016+679.+Arricchito+con+riferimenti+ai+Considerando+Aggiornato+alle+rettifiche+pubblicate+sulla+Gazzetta+Ufficiale++dell+%27Unione+europea+127+del+23+maggio+2018.
- Rohit Malpani, and About World Health Organization (WHO) More knowledge from World Health Organization (WHO). “Ethics and Governance of Artificial Intelligence for Health: Who Guidance.” The Communication Initiative Network, 2021, <https://www.comminit.com/global/content/ethics-and-governance-artificial-intelligence-health-who-guidance>
- Rohit Malpani, and About World Health Organization (WHO) More knowledge from World Health Organization (WHO). “Ethics and Governance of Artificial Intelligence for Health: Who Guidance.” The Communication Initiative Network, 2021, si veda <https://www.comminit.com/global/content/ethics-and-governance-artificial-intelligence-health-who-guidance>
- Roncaglia G. “Computer Che Copiano: Test Di Turing, Web Corpora e Filtraggio Collaborativo”, 2014, si veda <http://hdl.handle.net/2067/2614> .
- Ryan, Mark, and Bernd Carsten Stahl. "Artificial intelligence ethics guidelines for developers and users: clarifying their content and normative implications" in Journal of Information, Communication and Ethics in Society, 2020.
- Saint James Aquino, Yves, et al. "Utopia versus Dystopia: Professional perspectives on the impact of healthcare Artificial Intelligence on clinical roles and skills" in International Journal of Medical Informatics, 2022, p. 104903.
- Scalzini, S. "Alcune questioni a proposito di algoritmi, dati, etica e ricerca" in Rivista italiana di medicina legale e del diritto in campo sanitario, 2019, pp.169-178.
- .

- Schönberger D., “*Artificial intelligence in healthcare: a critical analysis of the legal and ethical implications*”, *International Journal of Law and Information Technology*, 2017 V. 27, pp. 171–203, <https://doi.org/10.1093/ijlit/eaz004>
- Traoré, Mariam, and Shuichiro Yamamoto. "Healthcare CloudEcosystem Risk Analysis and Modeling: A FAIR Approach—A Case Study of Arterys™ on AWS." 2018, 7th International Congress on Advanced Applied Informatics, DOI:10.1109/IIAI-AAI.2018.00171.
- V Dignum, ‘*The ART of AI—Accountability, Responsibility, Transparency, Medium*’, 2019, <https://me>
- Votano G., “*Intelligenza artificiale in ambito sanitario: il problema della responsabilità civile*”, in *Danno e Responsabilità*, 2022, n. 6 p. 669;
- "White Paper on Artificial Intelligence: a European approach to excellence and trust". European Commission, commission.europa.eu/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en.
- Wachter, Sandra, et al. "*Counterfactual Explanations Without Opening the Black Box: Automated Decisions and the GDPR*" *SSRN Electronic Journal*, 2017
- Wachter, Sandra, et al. "*Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation*". *SSRN Electronic Journal*, 2016.
- Wiener, N. "*Some Moral and Technical Consequences of Automation*" in *Science*, 1960 vol. 131, n. 3410, pp. 1355-58.
- Yüksel, A. E. B. “A General Look at the Proposal for a Regulation on Artificial Intelligence of the European Commission” *Türkiye Adalet Akademisi Dergisi*, 2022, pp. 19–46
- “*A Definition of Artificial Intelligence: Main Capabilities and Scientific Disciplines.*” *Shaping Europe's Digital Future*, <https://digital-strategy.ec.europa.eu/en/library/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines>;
-