

*T4F Series*  
PAPER N. 2  
a.a. 2022/2023

**Il progetto *Augmented Point of Care*,  
risvolti giuridici:  
la ricerca di un equilibrio di  
*engagement* tra professionista  
sanitario e sistema di Intelligenza  
Artificiale**

AURORA CALECA, VITTORIA DA ROS,  
ANNA FUSO, LINDA VARANZANO

Trento BiLaw Selected Student Papers

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

**Il progetto *Augmented Point of Care*, risvolti giuridici:  
la ricerca di un equilibrio di *engagement* tra professionista sanitario e sistema di  
Intelligenza Artificiale**

*Aurora Caleca, Vittoria Da Ros, Anna Fuso, Linda Varanzano\**

ABSTRACT: Taking as a reference the experience of the *Augmented Point of Care* project (AGIR), concerning the creation of an experimental point of care based on Artificial Intelligence (AI) and Augmented Reality technologies, the paper aims at analyzing the main legal issues due to the application of Artificial Intelligence within the medical field. In this way, the paper will investigate the functioning of the intakes and point of care phases in diagnosing procedures in order to verify their compatibility with the current legal framework, highlighting the potential but also the criticalities of the AGIR project.

KEYWORDS: AI; healthcare; medical devices; AI Act; GDPR

SOMMARIO: 1. Introduzione: il progetto AGIR Augmented Point of Care – 1.1. Progetto AGIR – 1.2. Accettazione e Point of Care: anamnesi e diagnosi – 2. Riflessioni giuridiche. Medical Device Regulation – 2.1. General Data Protection Regulation – 2.2. Artificial Intelligence Act – 2.3. Consenso informato – 3. Vantaggi e svantaggi – 3.1. Pro: accesso alla salute e riduzione margine di errore – 3.2. Contro: pregiudizio algoritmico, taglio delle code e conservative profiling – 4. Sostenibilità del progetto e prospettive future.

### **1. Introduzione: il progetto AGIR Augmented Point of Care**

Presso il Dipartimento di Ingegneria e Scienze dell'Informazione (DISI) dell'Università di Trento, un gruppo di ricerca diretto dal Professor Paolo Giorgini (head department del DISI) e il Professor Giuseppe Riccardi (referente del progetto) stanno lavorando al progetto AGIR.

Con il presente lavoro si intende approfondire il funzionamento della fase di accettazione e Point of Care, che gestiscono rispettivamente le procedure di anamnesi e diagnosi, al fine di verificarne la compatibilità con il quadro giuridico attuale, mettendo in luce potenzialità ma anche le criticità del progetto AGIR.

---

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

### 1.1 Progetto AGIR

Il progetto si sostanzia nella realizzazione di una sala per l'accettazione, un Point of Care e una sala operatoria; si tratta di spazi sia fisici che virtuali, con schermi e visori per la realtà aumentata, sviluppati con sistemi di intelligenza artificiale.

Il progetto si svolge in collaborazione con i medici dell'Ospedale Santa Chiara di Trento, i quali - nella fase di progettazione - affiancano i programmatori del DISI per la stesura dell'algoritmo e il suo successivo training. Inoltre, una volta completata la sala operatoria, questa sarà un banco di prova sia per i medici, ma anche per gli stessi studenti di medicina dell'Università di Trento, i quali avranno la possibilità di interfacciarsi con i sistemi di intelligenza artificiale al fine di esercitarsi sui vari interventi chirurgici che si svolgeranno nelle sale del Santa Chiara.

Come si coordinano l'attività umana e il software? La finalità del progetto è sperimentare la collaborazione tra uomo e macchina, ma non solo: sarà anche una interazione multi-entità, ciò significa che la cooperazione avverrà anche tra le varie macchine impiegate le quali si scambieranno dati e, inoltre, saranno capaci di controllare il lavoro delle altre macchine.

È bene ora illustrare la composizione delle varie sale.

La fase di accettazione ed anamnesi sarà gestita in autonomia da un robot umanoide ARI<sup>1</sup>, attualmente in fase sperimentale presso le aule del DISI, ma si prevede l'assistenza umana qualora tali software saranno utilizzati negli ospedali. Ne segue lo stesso iter il Point of Care in cui, allo stadio sperimentale, gli esami preoperatori (basti pensare ad ecografie, radiografie, etc.) verranno svolti autonomamente da sistemi AI3, ma in ottica di una futura applicazione pratica, vedrà l'assistenza di un professionista sanitario.

La sala operatoria, in fase di ricerca, sarà dotata di telecamere e vari sensori al fine di monitorare come i professionisti sanitari riescono ad interagire con la macchina.

I dati raccolti implementano il sistema stesso, infatti, i sistemi di intelligenza artificiale utilizzati nel progetto AGIR sono machine learning<sup>2</sup>, ciò significa che lo stesso software imparerà dai medici che si eserciteranno nella sala operatoria, dai dati raccolti nella fase di accettazione e anche dagli esami svolti nel Point of Care.

---

<sup>1</sup> Durante il progetto SPRING, PAL Robotics ha progettato e prodotto ARI i cui compiti nell'ambiente ospedaliero comprenderanno l'accoglienza dei pazienti e visitatori nella sala d'attesa, l'assistenza ai moduli di accettazione e dimissioni, l'assistenza agli appuntamenti medici, l'indirizzamento dei pazienti ai loro appuntamenti e l'intrattenimento dei visitatori. Per approfondire si consultino *ARI and the SPRING Project, ARI social robots delivered to partners to prepare for hospital use in SPRING Project*, in *PAL Robotics*, 23 dicembre 2021, in <https://blog.pal-robotics.com/ari-social-robot-delivered-to-hospitals/#:~:text=The%20SPRING%20Project%20has%20reached,deployment%20in%20a%20hospital%20environment> (ultimo consultazione 5/10/2023); M. HAMPSON, *PAL Robotics customizable bots could be the next frontline workers, The robots could interact with COVID-19 patients, check their temperature, and even provide entertainment*, in *IEEE Spectrum*, 20 maggio 2020, in <https://spectrum.ieee.org/pal-robotics-customizable-bots-could-be-the-next-frontline-workers> (ultimo consultazione 5/10/2023).

<sup>2</sup> Questa tecnica di apprendimento viene definita «*Branca dell'AI che si occupa dello sviluppo di algoritmi e tecniche finalizzate all'apprendimento automatico mediante la statistica computazionale e l'ottimizzazione matematica*». In *Machine learning*, in

Un progetto di questo genere ci pone di fronte vari interrogativi a cui forse la nostra società non è ancora pronta a rispondere? Arriverà un momento in cui preferiremo che sia un robot a curarci, oppure ci fideremo sempre e solo del medico umano? Quanto il singolo è disposto a fidarsi di un sistema di intelligenza artificiale e ad affidargli la sua salute?

Con AGIR si vuole tentare di dare le prime risposte a queste domande che spaventano, ad oggi, la maggior parte delle persone.

## 1.2 Accettazione e Point of Care: anamnesi e diagnosi

La prima fase di anamnesi consisterà nella raccolta, tramite la voce del paziente, delle informazioni che potranno aiutare per l'elaborazione di una corretta diagnosi. La raccolta avverrà tramite i sensori presenti nell'umanoide ARI che interagirà fin da subito col paziente.

Successivamente verranno svolti l'esame obiettivo e gli esami strumentali da parte dei robot presenti nel Point of Care. Il primo consiste in una valutazione obiettiva e oggettiva delle condizioni del paziente. I secondi consistono negli esami necessari per analizzare la condizione fisica del corpo. Tramite questi tre passaggi si arriverà ad ottenere la diagnosi. Questa sarà suggerita dal sistema di intelligenza artificiale, ma avrà bisogno di una conferma da parte del professionista sanitario.

Le fasi in questione sono fondamentali per poter fare un confronto con altri sistemi di intelligenza artificiale che sono stati creati per svolgere questa funzione. Oltre all'interazione diretta con un umanoide, come avverrà nel progetto AGIR, a livello globale sono state create anche piattaforme alle quali l'utente può connettersi per ricevere assistenza medica immediata oppure i sistemi di intelligenza artificiale sono stati utilizzati come supporto alla diagnostica permettendogli di analizzare direttamente le immagini radiologiche e di classificare i dati ottenuti dalla loro analisi<sup>3</sup>.

Il progetto AGIR è innovativo e si distanzia da altri sistemi come Babylon<sup>4</sup> - fornitore di servizi sanitari attivo in Regno Unito che offre consulenze a distanza con operatori sanitari - perché non interviene in aiuto del medico, ma prevede che il sistema di intelligenza artificiale esegua l'attività medica in autonomia e che necessiti solo di una conferma da parte del professionista sanitario. Il progetto AGIR opera nel concreto in diretto contatto con il paziente, ciò è fondamentale perché il contesto influisce fortemente sull'operatività del sistema: la stessa definizione di sistema di intelligenza artificiale «*a machine-based system that is*

---

Treccani.it - Vocabolario Treccani on line, 2019, in [https://www.treccani.it/vocabolario/machine-learning\\_%28Neologismi%29/](https://www.treccani.it/vocabolario/machine-learning_%28Neologismi%29/) (ultima consultazione 18/09/2023).

<sup>3</sup> Per approfondire G. PASCERI, C. BIBBOLINO, M. STASI, M. FORACCHIA, G. PIROTTA, F. COPPOLA, E. NERI, R. GRASSI (DOCUMENTO SIRM), *Intelligenza artificiale in radiologia*, Milano, 2020, in <https://sirm.org/wp-content/uploads/2021/04/317-Documento-SIRM-2020-Intelligenza-artificiale-in-radiologia.pdf> (ultima consultazione 5/10/2023).

<sup>4</sup> Per approfondire C. CASONATO, *Intelligenza artificiale e medicina: l'impatto sulla relazione di cura (cenni)*, in U. SALINASTRO (a cura di), *SMART la persona e l'infosfera*, Pisa, 2022, p. 107.

*designed to operate with varying levels of autonomy and that can, for explicit or implicit objectives, generate outputs such as predictions, recommendations, or decisions that influence physical or virtual environments»<sup>5</sup>*  
dà rilevanza sia alla competenza di *perceiving*, cioè la capacità di cogliere alcune caratteristiche dell'ambiente in cui è inserita<sup>6</sup>, sia di *deciding* che le permette di modificare in maniera sostanziale l'ambiente che la circonda.

In seguito all'anamnesi e alla diagnosi, vi sarà una terza fase chirurgica che porterà invece a una vera e propria collaborazione uomo-macchina, almeno nella fase iniziale, per far sì che i sistemi di intelligenza artificiale coinvolti possano apprendere dall'esperienza umana. Non tratteremo nello specifico l'analisi dello stadio chirurgico in quanto solo in un futuro lontano si potrà arrivare a far operare in autonomia le macchine direttamente sul paziente senza coinvolgimento umano. Infatti, per il momento è stato applicato sugli animali, in particolare un maiale, con il metodo STAR<sup>7</sup>.

## 2. Riflessioni giuridiche. Medical Device Regulation

Il recente Reg. UE 2017/745 sui dispositivi medici -MDR- impatta fortemente nell'ambito dei software di area medica.

Il regolamento modifica la direttiva 2001/83/CE, il regolamento (CE) n. 1223/2009, il regolamento (CE) n.178/2002; vengono poi abrogate le direttive 90/385/CEE e 93/42/CEE del Consiglio.

La disciplina si applica a decorrere dal 26 maggio 2020; inoltre entro il 27 maggio 2027, la Commissione valuterà l'applicazione del presente regolamento ed elaborerà una relazione di valutazione dei progressi compiuti nella realizzazione degli obiettivi ivi contenuti, compresa una valutazione delle risorse necessarie per l'attuazione della disciplina.

Il regolamento stabilisce le norme relative all'immissione sul mercato, la messa a disposizione sul mercato o la messa in servizio dei dispositivi medici per uso umano e degli accessori per tali dispositivi nell'Unione. Si applica inoltre alle indagini cliniche relative a tali dispositivi medici e relativi accessori condotte nell'Unione.

---

<sup>5</sup> Art.3, proposta di Regolamento del Parlamento europeo e del Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale (legge sull'intelligenza artificiale) e modifica alcuni atti legislativi dell'Unione.

<sup>6</sup> C. CASONATO, *Intelligenza artificiale e giustizia: potenzialità e rischi*, in *DPCE online*, 3, 2020, pp. 3369-3389.

<sup>7</sup> H. SAEIDI, J. D. OPFERMANN, M. KAM, S. WEI, S. LEONARD, M. H. HSIEH, J.U. KANG, A. KRIEGER, *Autonomous robotic laparoscopic surgery for intestinal anastomosis*, in *Science Robotics*, 62, 2022, 1-13. STAR (Smart Tissue Autonomous Robot) è un robot che ha eseguito in completa autonomia un intervento di chirurgia addominale in laparoscopia. Questo sistema di intelligenza artificiale è stato progettato dalla John Hopkins University e dal Children's National Hospital di Washington negli Stati Uniti. È un sistema a guida visiva che è stato progettato dall'ingegnere Axel Krieger che commenta: «*What makes the STAR special is that it is the first robotic system to plan, adapt, and execute a surgical plan in soft tissue with minimal human intervention*». La creazione di questo sistema competente ad eseguire l'anastomosi viene considerato come un passo importante per il settore chirurgico.

Il presente regolamento istituisce inoltre la banca dati europea sui dispositivi medici Eudamed che migliorerà la trasparenza e la condivisione delle informazioni sui dispositivi medici disponibili sul mercato UE, come evidenziato al considerando n. 41.

All'art. 2, n. 1) viene data la seguente definizione di dispositivo medico: «qualunque strumento, apparecchio, apparecchiatura, software, impianto, reagente, materiale o altro articolo, destinato dal fabbricante a essere impiegato sull'uomo, da solo o in combinazione, per una o più delle seguenti destinazioni d'uso mediche specifiche: diagnosi, prevenzione, monitoraggio, previsione, prognosi, trattamento o attenuazione di malattie; diagnosi, monitoraggio, trattamento, attenuazione o compensazione di una lesione o di una disabilità; studio, sostituzione o modifica dell'anatomia oppure di un processo o stato fisiologico o patologico; fornire informazioni attraverso l'esame in vitro di campioni provenienti dal corpo umano, inclusi sangue e tessuti donati, e che non esercita nel o sul corpo umano l'azione principale cui è destinato mediante mezzi farmacologici, immunologici o metabolici, ma la cui funzione può essere coadiuvata da tali mezzi».

La Regola 11 dell'Allegato VIII, allegato riguardante la classificazione dei dispositivi, stabilisce che: «Il software destinato a fornire informazioni utilizzate per prendere decisioni a fini diagnostici o terapeutici rientra nella classe IIa, a meno che tali decisioni abbiano effetti tali da poter causare: il decesso o un deterioramento irreversibile delle condizioni di salute di una persona, nel qual caso rientra nella classe III, o un grave deterioramento delle condizioni di salute di una persona o un intervento chirurgico, nel qual caso rientra nella classe IIb. Il software destinato a monitorare i processi fisiologici rientra nella classe IIa, a meno che sia destinato a monitorare i parametri fisiologici vitali, ove la natura delle variazioni di detti parametri sia tale da poter creare un pericolo immediato per il paziente, nel qual caso rientra nella classe IIb. Tutti gli altri software rientrano nella classe I».

Dalla lettura di queste norme si comprende come tutti i software che hanno finalità di diagnosi e cura e che supportano l'operatore sanitario ad assumere una decisione terapeutica oppure aiutano l'erogazione della prestazione stessa devono rientrare nella nozione di dispositivo medico.

Il Capo II tratta della messa a disposizione sul mercato e messa in servizio dei dispositivi. Vengono quindi ribaditi gli obblighi generali dei fabbricanti, i quali devono garantire che i dispositivi siano stati progettati e fabbricati conformemente al regolamento. Si stabilisce inoltre che i fabbricanti, all'interno della loro organizzazione, dispongono di almeno una persona responsabile del rispetto della normativa che possieda le competenze necessarie nel settore dei dispositivi medici<sup>8</sup>.

Il regolamento si occupa anche di disciplinare il settore della valutazione clinica e delle indagini cliniche, a ciò è deputato il Capo VI.

---

<sup>8</sup> Artt. 10 e 15, Reg. (UE) 2017/745.

Inoltre, ai fini del presente lavoro è bene prendere in considerazione gli Allegati I e VIII.

All'Allegato I sono disciplinati i requisiti generali di sicurezza e prestazione. In relazione al progetto AGIR, in particolare, è opportuno richiamare l'art.14 (relativo all'interazione dei dispositivi con l'ambiente), l'art.15 (relativo ai requisiti di progettazione e fabbricazione) e all'art.17 (relativo ai software che costituiscono dispositivi a sé stanti). L'Allegato VIII è dedicato ai criteri di classificazioni.

In particolare, i dispositivi medici sono raggruppati, in funzione della loro complessità e del potenziale rischio per il paziente, in quattro classi: I- IIa-IIb- III.

I software impiegati nel progetto AGIR sono dispositivi a lungo termine, in quanto destinati a essere utilizzati per una durata continua superiore a 30 giorni.

La regola di applicazione 3.3 dell'Allegato VIII afferma che il software destinato a far funzionare un dispositivo o a influenzarne l'uso rientra nella stessa classe del dispositivo. Se il software non è connesso con nessun altro dispositivo è classificato separatamente.

Viene in rilievo quindi la Regola 11 (il cui testo è già stato riportato sopra) secondo la quale i software rientrano nella classe IIa qualora abbiano fini diagnostici o terapeutici o siano destinati a monitorare i processi fisiologici; rientrano invece nella classe III quando le decisioni prese dal software abbiano effetti tali da poter causare il decesso o il deterioramento irreversibile delle condizioni di salute. Per quest'ultima ipotesi il software potrebbe anche appartenere alla classe IIb. Tutti gli altri software rientrano nella classe I.

Sulla nuova classificazione dei software è stata poi emanata una specifica linea guida dal Medical Device Coordination Group<sup>9</sup>: la MDCG 2019-11 Guidance on Qualification and Classification of Software in Regulation (UE) 2017/745.

Sono state in seguito pubblicate altre due linee guida: MDCG 2019-16 Guidance on Cybersecurity for medical devices<sup>10</sup> e MDCG 2020-1 Guidance on Clinical Evaluation (MDR)/Performance Evaluation (IVDR) of Medical Device Software<sup>11</sup>.

Nel primo documento menzionato viene data la definizione di Medical Device Software (MDSW): «*Medical device software is software that is intended to be used, alone or in combination, for a purpose as specified in the definition of a "medical device" in the MDR*».

---

<sup>9</sup> Il gruppo di coordinamento dei dispositivi medici si occupa di questioni chiave nel settore dei dispositivi medici, dalla supervisione degli organismi notificati o dalla standardizzazione della sorveglianza del mercato, passando per questioni internazionali, nuove tecnologie e indagini cliniche. La sua competenza deriva dalla suddivisione in 13 sottogruppi, che rispettivamente forniscono consulenza e redigono linee guida sul proprio settore di competenza. I membri dei sottogruppi sono nominati dagli Stati Membri per un periodo di tre anni. I sottogruppi si riuniscono regolarmente con la presidenza della Commissione UE.

<sup>10</sup> MEDICAL DEVICE COORDINATION GROUP, *MDCG 2019-16 Guidance on Cybersecurity for medical devices*, dicembre 2019, <https://ec.europa.eu/docsroom/documents/41863> (ultima consultazione 5/10/2023).

<sup>11</sup> MEDICAL DEVICE COORDINATION GROUP, *MDCG 2020-1 Guidance on Clinical Evaluation (MDR)/Performance Evaluation (IVDR) of Medical Device Software*, marzo 2020, <https://ec.europa.eu/docsroom/documents/40323> (ultima consultazione 5/10/2023).



È importante chiarire che non tutti i software utilizzati nell'ambito dell'assistenza sanitaria sono qualificati come dispositivi medici. Un software che ha lo scopo di elaborare, creare o modificare informazioni mediche può essere qualificato come un MDSW se la creazione o la modifica di quella informazione è accompagnato da una finalità medica.

Bisogna ora chiedersi se, alla luce della precedente analisi, il robot ARI e i macchinari che saranno utilizzati nel Point of Care possano dirsi dispositivi medici tenendo in considerazione anche la finalità del progetto AGIR.

Il robot ARI potrebbe rientrare tra i c.d. «Hospital Information System» i quali vengono utilizzati per l'accettazione, gestione degli appuntamenti dei pazienti, per scopi assicurativi e di fatturazione. Questo tipo di sistemi non sono qualificati come dispositivi medici.

Per come, ad oggi, il robot ARI è stato programmato e per le funzioni che andrà a svolgere nelle prime fasi del progetto AGIR - che si svolgeranno nelle aule del DISI- non può farsi rientrare nella definizione di dispositivo medico, di conseguenza non dovranno essere rispettate le norme del MDR.

Nulla esclude che in futuro possano essere implementate le funzioni del robot ARI, arrivando anche ad affiancare il medico (o forse a sostituirlo) nel fornire l'anamnesi e diagnosi. Nel momento in cui ciò diventerà realtà, la disciplina del MDR troverà applicazione; per quanto concerne la classificazione del software come dispositivo medico, questo rientrerà nella classe IIa.

I macchinari impiegati nel Point of Care saranno in grado di eseguire in autonomia, ma è comunque prevista la cooperazione con il medico, esami e test preoperatori quali ECG, ecografie e radiografie. In questo caso si coglie direttamente la finalità medica del software e quindi l'applicabilità del MDR. Anche tali dispositivi faranno parte della classe IIa.

## 2.1 General Data Protection Regulation

Il Regolamento (UE) 2016/679 relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati, noto anche come Regolamento generale sulla protezione dei dati o GDPR (General Data Protection Regulation), è la principale normativa europea in materia di protezione dei dati personali. Ha l'obiettivo di armonizzare la protezione dei dati personali all'interno degli Stati Membri dell'Unione europea. Tra i vari diritti e libertà riconosciuti dal regolamento sono rilevanti ai fini del progetto in analisi le definizioni riguardanti i dati, la disciplina riguardante il trattamento dei dati personali e il consenso.

In particolare, il Regolamento definisce all'art. 4 n. 14-15 due categorie di dati. La prima riguarda i «dati biometrici» definiti come i «dati personali ottenuti da un trattamento tecnico specifico relativi alle caratteristiche fisiche, fisiologiche o comportamentali di una persona fisica che ne consentono o confermano l'identificazione univoca, quali l'immagine facciale o i dati dattiloscopici». La seconda riguarda i dati relativi alla salute cioè «i dati personali attinenti alla salute fisica o mentale di una persona fisica, compresa la prestazione di servizi di assistenza sanitaria, che rivelano informazioni relative al suo stato di salute». Per interpretare meglio queste definizioni è rilevante il Considerando 35 GDPR che prevede: «Nei dati personali relativi alla salute dovrebbero rientrare tutti i dati riguardanti lo stato di salute dell'interessato che rivelino informazioni connesse allo stato di salute fisica o mentale passata, presente o futura dello stesso. Questi comprendono [...] qualsiasi informazione riguardante, ad esempio, una malattia, una disabilità, il rischio di malattie, l'anamnesi medica, i trattamenti clinici o lo stato fisiologico o biomedico dell'interessato».

All'interno del regolamento queste categorie vengono definite come «soggette a trattamento speciale» perché è necessario il consenso esplicito dell'interessato ed è presente un vincolo per quanto riguarda le finalità (Art. 9 GDPR).

Il Regolamento adotta un approccio basato sulla valutazione del rischio con il quale si determina la misura di responsabilità del titolare o del responsabile del trattamento che devono tutelare i dati dell'interessato rispettando determinate finalità.

Il trattamento dei dati viene disciplinato in particolare al comma 1 dell'art.24 GDPR che impone al titolare l'obbligo di mettere in atto misure tecniche e organizzative adeguate a garantire ed essere in grado di dimostrare che il trattamento è effettuato conformemente al Regolamento. Il titolare ha anche il dovere di trattare i dati dell'interessato rispettando le finalità per le quali ha dato il consenso in modo che il trattamento sia considerato lecito (art. 6 GDPR).

Le disposizioni analizzate sono rilevanti ai fini del caso pratico poiché il paziente deve interagire direttamente con un umanoide che ha il compito di accoglierlo e di raccogliere i suoi dati. In questo modo la macchina può

avere accesso per intero al suo fascicolo sanitario. Successivamente prende nota dei sintomi elencati dal paziente che verranno tradotti in dati. Questi vengono raccolti dal sistema per poi essere utilizzati a fini di ricerca e di apprendimento per i dispositivi.

Per eseguire queste operazioni l'umanoide utilizza videocamere, microfoni e sensori tramite i quali registra quanto detto e analizza l'interazione tra uomo e macchina.

Nella prima fase di apprendimento la macchina collabora con il medico raggiungendo un punto di equilibrio, ovvero una situazione di uguale lavoro tra il professionista sanitario e il sistema. Si prevede la possibilità che in futuro la macchina potrà svolgere il tutto in autonomia.

Sorge anche un problema riguardo il trattamento dei dati in quanto i sistemi di AI presentano il problema della black box<sup>12</sup>, cioè la macchina elabora i dati senza spiegare i passaggi che le hanno permesso di raggiungere l'output. Questo crea un problema riguardo al principio del rispetto delle finalità per le quali si mettono a disposizione i dati personali dell'interessato poiché la macchina potrebbe usarli anche per finalità non determinate precedentemente e porta ad una possibile violazione dell'obbligo di trasparenza. Inoltre, in questa situazione risulta difficile individuare il titolare<sup>13</sup>, ovvero colui che determina le finalità e i mezzi del trattamento, e il responsabile<sup>14</sup> del trattamento, cioè la persona fisica o giuridica che tratta i dati personali per conto del titolare, poiché i dati vengono elaborati dalla macchina, ma è necessario capire se essa può ricoprire questi ruoli.

In futuro sarà inoltre possibile far svolgere alla macchina il processo di formazione della diagnosi che le permetterà di prendere una decisione in modo automatizzato e questo porta ad una possibile violazione dell'art. 22 GDPR che riconosce in capo all'interessato «il diritto di non essere sottoposto a una decisione basata unicamente sul trattamento automatizzato, compresa la profilazione, che produca effetti giuridici che lo riguardano o che incida in modo analogo significativamente sulla sua persona». Questo diritto però non può essere rivendicato nel caso in cui l'interessato abbia prestato il consenso (art. 22, lettera c).

Si può quindi parlare dell'approccio Human-in-the-loop che riguarda la capacità dell'uomo di essere coinvolto negli automatismi e di perfezionare i modelli di Machine Learning. L'umano ha un ruolo di «mediazione»: infatti nelle prime fasi di creazione di sviluppo del progetto AGIR è previsto un contributo umano da parte di medici esperti che renderà lecita l'elaborazione dei dati e la futura presa di decisione da parte del sistema

---

<sup>12</sup> C. CASONATO, *Intelligenza artificiale e giustizia: potenzialità e rischi*, cit., pp. 3369-3389.

<sup>13</sup> Art. 4, par.1, n. 7), Reg. (UE) 2016/679: «titolare del trattamento: 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; quando le finalità e i mezzi di tale trattamento sono determinati dal diritto dell'Unione o degli Stati membri, il titolare del trattamento o i criteri specifici applicabili alla sua designazione possono essere stabiliti dal diritto dell'Unione o degli Stati membri».

<sup>14</sup> Art. 4, par.1, n. 8), Reg. (UE) 2016/679: «responsabile del trattamento: la persona fisica o giuridica, l'autorità pubblica, il servizio o altro organismo che tratta dati personali per conto del titolare del trattamento».

poiché non sarà interamente automatizzata. In un futuro prossimo potrà esserci la possibilità di sottoporre l'umano ad una prestazione interamente automatizzata e quindi si potrà applicare l'eccezione del consenso. Un altro profilo importante è quello riguardante l'art. 25 GDPR che introduce il concetto di «*privacy by design*», cioè della protezione dei dati fin dalla progettazione. La tutela dei dati personali degli utenti dev'essere integrata e presente durante il ciclo di progettazione del sistema e anche di esecuzione del trattamento stesso. È necessario trovare un punto d'equilibrio nell'interazione tra uomo e macchina attraverso l'applicazione di questo principio.

## 2.2 Artificial Intelligence Act

L'Artificial Intelligence Act (da qui in poi AI Act) è una proposta di Regolamento dell'Unione europea in materia di controllo dell'intelligenza artificiale<sup>15</sup>. La discussione per l'adozione di tale documento iniziò nell'aprile 2021 con una proposta della Commissione Europea. Esso diverrà legge in seguito al buon esito di una negoziazione tra Parlamento europeo, Consiglio dell'Unione europea e Commissione europea.

L'approccio adottato dall'AI Act per la regolamentazione dei sistemi d'intelligenza artificiale comprende la «piramide del rischio». In base al rischio insito negli ambiti di applicazione delle intelligenze artificiali da regolamentare vi sono infatti diversi obblighi e requisiti, andando dai sistemi proibiti passando poi per i sistemi ad alto rischio, basso rischio per poi finire con quelli a rischio minimo. Le regole per la classificazione di sistemi ad alto rischio si trovano nell'articolo 6 della proposta. Questa, nel primo paragrafo, spiega come siano classificabili come sistemi ad alto rischio tutti quei sistemi che sono un prodotto o la componente di un prodotto, il quale è disciplinato dalla normativa di armonizzazione elencata nell'allegato II. Andando a vedere l'elenco delle norme all'interno del suddetto allegato notiamo come vi sia anche il Medical Device Regulation sopra citato. Di conseguenza, già dal primo paragrafo dell'articolo possiamo notare come il progetto AGIR rientrerebbe nella classificazione di sistema ad alto rischio. Nel secondo paragrafo dello stesso articolo, però, notiamo come vi siano ulteriori parametri che rendono un sistema di intelligenza artificiale qualificabile come ad alto rischio. A questo proposito il rimando è all'allegato III. Nel suddetto allegato notiamo come siano classificati come sistemi ad alto rischio i sistemi di: «1. Identificazione e categorizzazione biometrica delle persone fisiche: a) i sistemi di IA destinati a essere utilizzati per l'identificazione biometrica remota in tempo reale e a posteriori delle persone fisiche. (...) 5. Accesso a prestazioni e servizi pubblici e a servizi privati essenziali e fruizione degli stessi: Accesso a prestazioni e servizi pubblici e a servizi privati essenziali e fruizione

---

<sup>15</sup> Proposta di Regolamento del Parlamento europeo e del Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale (legge sull'intelligenza artificiale) e modifica alcuni atti legislativi dell'Unione.

degli stessi: a) i sistemi di IA destinati a essere utilizzati dalle autorità pubbliche o per conto di autorità pubbliche per valutare l'ammissibilità delle persone fisiche alle prestazioni e ai servizi di assistenza pubblica, nonché per concedere, ridurre, revocare o recuperare tali prestazioni e servizi»<sup>16</sup>.

La tecnologia del progetto AGIR lavorerà anche tramite l'identificazione e la categorizzazione dei dati biometrici visionando, ed in seguito costruendo, l'anamnesi delle persone che si trova di fronte. Possiamo inoltre vedere come nel punto di accettazione la macchina assegnerà la priorità con la quale una persona potrà accedere ai servizi ospedalieri, come quelli di pronto soccorso, tramite l'attribuzione dei c.d. «codici». A ben vedere anche tramite le eventuali analisi fatte nel Point of Care la macchina potrebbe decidere eventualmente di tenere sotto osservazione una persona piuttosto che sulla priorità di un'operazione.

I sistemi ad alto rischio sono soggetti ad una serie di requisiti applicabili al fornitore (o agli eventuali importatori), i distributori, così come degli utenti. Tra questi requisiti grande importanza è data alla supervisione umana. Come già illustrato, gli ingegneri puntano a trovare l'equilibrio perfetto tra l'umano e la macchina, magari anche con l'obiettivo futuro di rendere la macchina completamente autonoma.

Sempre nell'introduzione, ci siamo interrogate su un possibile scenario in cui le persone siano disponibili a farsi curare da un robot o da un medico umano, notando come la società non sia ancora pronta a dare una risposta ad un tale interrogativo. Un panorama simile risulta, per il momento, lontano. L'articolo 14 dell'AI Act richiede, infatti, che i sistemi IA ad alto rischio siano progettati con un'interfaccia uomo-macchina adeguata ad un efficace controllo da parte di un agente umano. Si può notare come l'approccio utilizzato dall'AI Act rimanga piuttosto umano-centrico<sup>17</sup>. E nello stesso articolo viene anche il motivo della scelta di imporre una sorveglianza umana: proprio tramite la sorveglianza umana si punta a ridurre al minimo i rischi per la salute, la sicurezza o i diritti fondamentali che possono emergere dall'uso della tecnologia. I principali rischi che potrebbero emergere dall'uso delle macchine nell'ambito medico (in completa assenza di una sorveglianza umana) sono approfonditi più avanti (par. 2.4).

Alla luce della normativa vigente e futura un'illimitata autonomia decisionale rimane inattuabile (si veda l'art. 22 del GDPR, sopra riportato) e con grande probabilità tale linea verrà adottata anche in futuro, o almeno, sarà tale finché progetti simili ad AGIR non aiuteranno a darci risposte esaustive sulle prospettive di applicazione autonoma dei sistemi di intelligenza artificiale.

Il principio cardine dell'AI Act è di regolamentare l'IA affinché possa definirsi affidabile. Il concetto è particolarmente rilevante in ambito sanitario, considerando il delicatissimo ruolo che svolge e l'importanza di una collaborazione efficace tra medico e paziente. Si può notare come la caratteristica tipica dei sistemi di

---

<sup>16</sup> Allegato III, proposta di Regolamento del Parlamento europeo e del Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale (legge sull'intelligenza artificiale) e modifica alcuni atti legislativi dell'Unione.

<sup>17</sup> O. KUDINA, *Regulating AI in Health Care: The Challenges of Informed User Engagement*, in *Hastings Center Report*, 5, 2021, pp. 6-7.

machine learning - e cioè il fenomeno Black Box- abbia portato problematiche in merito non indifferenti. In merito abbiamo avuto la possibilità di confrontarci con il professore G. Riccardi, in cui è emersa l'importanza di avere una tecnologia che dia le sue risposte corredate dalla percentuale di esattezza da questa calcolata. Il punto riteniamo sia cruciale ed è in linea anche con il requisito di trasparenza descritto nell'articolo 13 della proposta di regolamento. Altro aspetto rilevante ai fini dell'applicazione dei sistemi di intelligenza artificiale in ambito sanitario è la caratteristica del *sapere di non sapere*. In proposito, il professore G. Riccardi ha espresso la rilevanza della *rejection*: ciò che potremmo definire come una regolazione *by design* in cui la macchina lasci autonomamente il controllo all'umano perché arrivata alla consapevolezza di non essere in grado di prendere una decisione o di non averne i mezzi idonei. Da una necessità pratica, risulta che, attualmente, l'intervento umano rimanga di grande importanza sia per quanto concerne la scelta di seguire un suggerimento dato dalla macchina (anche tenendo in conto la percentuale mostrata dalla stessa) ma anche in caso in cui la macchina decida di arrestarsi in determinate situazioni.

A questo punto ci si potrebbe chiedere in che modo costruire efficacemente la fiducia di cui tanto necessitiamo. La linea della Commissione europea sembra concentrarsi su un'interpretazione piuttosto ristretta del concetto di affidabilità, concentrandosi sull'accettabilità dei rischi derivati dall'uso del sistema di intelligenza artificiale<sup>18</sup>. Seguendo questa linea, infatti, è imposto un complesso sistema della gestione dei rischi «costituito da un processo iterativo continuo eseguito nel corso dell'intero ciclo di vita di un sistema di AI ad alto rischio, che richiede un aggiornamento costante e sistematico»<sup>19</sup>. Tale sistema di gestione dei rischi è basato sull'identificazione rischi noti e prevedibili associati a ciascun sistema ad alto rischio, la valutazione dei rischi che potrebbero emergere dall'uso del sistema nel suo ambito di applicazione (tenendo anche conto di un eventuale uso improprio ragionevolmente prevedibile), la valutazione di eventuali rischi emersi in seguito alla messa sul mercato e l'adozione di adeguate misure di eliminazione o quantomeno mitigazione dei rischi. Si può intuire nell'intero articolo l'importanza del dialogo fra le varie parti, dialogo già suggerito in dottrina come mezzo ideale per regolamentare l'intelligenza artificiale. Vi è infatti un ruolo chiave dato alle istruzioni fornite agli utenti, le quali devono essere perfettamente intelleggibili per gli stessi, così come al concetto di «un uso improprio ragionevolmente prevedibile», concetto che richiede una concretizzazione nell'ambito di applicazione, tenendo presente anche il tipo di utenti dal quale verrà utilizzato. Anche a questo

---

<sup>18</sup> J. LAUX, S. WACHTER, B. MITTELSTADT, *Trustworthy artificial intelligence and the European Union AI Act: On the Conflation of Trustworthiness and the Acceptability of Risk*, in *Regulation and Governance*, 2023, 1-30.

<sup>19</sup> Art. 9, par. 1, proposta di Regolamento del Parlamento europeo e del Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale (legge sull'intelligenza artificiale) e modifica alcuni atti legislativi dell'Unione.

proposito il progetto AGIR risulta essere piuttosto interessante dato che si prefigge di essere sperimentato con un contributo diretto del personale dell’Ospedale Santa Chiara.

L’ultimo dato da prendere in considerazione è l’importanza dei dati scelti per il training della macchina. Secondo l’attuale stato dell’AI Act, i dati devono essere «pertinenti, rappresentativi, esenti da errori e completi». Devono, inoltre, possedere «le proprietà statistiche appropriate, anche, ove applicabile, per quanto riguarda le persone o i gruppi di persone sui quali il sistema di intelligenza artificiale ad alto rischio è destinato a essere usato»<sup>20</sup>. Gli ingegneri del DISI hanno comunicato l’idea di utilizzare sistemi pre-addestrati nella fase iniziale per poi procedere con l’inserimento di dati raccolti dalle varie simulazioni che si svolgeranno in laboratorio. In seguito, non escludono la possibilità di un addestramento anche da dati forniti direttamente dall’ospedale coinvolto. Rileva comunque il fatto che i requisiti imposti dall’AI Act, al momento, risultano piuttosto difficili da concretizzare (ed infatti son stati già proposti degli emendamenti sia dal Parlamento che dal Consiglio europei) e il progetto AGIR non fa eccezione: innanzitutto gli ingegneri dovranno scegliere un sistema addestrato su dati che potrebbero essere validi anche per la popolazione alla quale sarà diretto il servizio e poi controllare continuamente che i dati inseriti rispettino lo stesso standard, cioè che siano, secondo l’attuale formulazione dell’AI Act, «dati totalmente rappresentativi, esenti da errori e completi». I sopracitati emendamenti potrebbero esserne d’aiuto, dato che mutano la prospettiva: da assolutista a relativista, in quanto mirano a stabilire la validità in ottica di ciò che è possibile (il Consiglio parla di «*to the best extent possible*») con un occhio al punto in cui è arrivato lo stato dell’arte (come proposto dal Parlamento). Si noti come queste proposte di emendamento stabiliscano comunque uno standard piuttosto elevato ma, al tempo stesso, più facilmente raggiungibile.

### 2.3 Consenso informato

I sistemi di intelligenza artificiale, in particolare i più sofisticati, sono connotati da *bias* di trasparenza che ostacolano una chiara e comprensibile comunicazione delle informazioni. Il *diritto di conoscere* degli utenti non contempla solo essere informati sulla natura del proprio interlocutore<sup>21</sup>, ma anche sapere le modalità di funzionamento del sistema con cui ci si rapporta. Questi sono solo alcuni dei motivi per cui l’ultima riflessione giuridica è dedicata alla disciplina del consenso informato.

Il consenso informato, sancito per la prima volta dal codice di Norimberga del 1946, è il principio in ambito sanitario secondo cui «*un intervento nel campo della salute non può essere effettuato se non dopo che la*

<sup>20</sup> Art. 10, proposta di Regolamento del Parlamento europeo e del Consiglio che stabilisce regole armonizzate sull’intelligenza artificiale (legge sull’intelligenza artificiale) e modifica alcuni atti legislativi dell’Unione.

<sup>21</sup> Per approfondire C. CASONATO, *Unlocking the Synergy: Artificial Intelligence and (old and new) Human Rights*, in *BioLaw Journal – Rivista di BioDiritto*, 3, 2023, pp. 233-240.

*persona interessata abbia dato consenso libero e informato*»<sup>22</sup> come dispone la Convenzione di Oviedo all'art. 5.

In Italia è stato disciplinato dalla legge 219/2017 in cui allo stesso art. 1 si prevede «(...) nessun trattamento sanitario può essere iniziato o proseguito se privo del consenso libero e informato della persona interessata, tranne che nei casi espressamente previsti dalla legge»<sup>23</sup>

Il consenso informato è oggi un diritto ma presenta criticità - pratiche, giuridiche ed etiche - non indifferenti e la questione si amplifica nell'ambito medico che coinvolge sistemi di intelligenza artificiale.

La Loi de Bioéthique 2021/1017 è stata la prima normativa che ne ha preso atto introducendo l'art. 17:

*«I.-Le professionnel de santé qui décide d'utiliser, pour un acte de prévention, de diagnostic ou de soin, un dispositif médical comportant un traitement de données algorithmique dont l'apprentissage a été réalisé à partir de données massives s'assure que la personne concernée en a été informée et qu'elle est, le cas échéant, avertie de l'interprétation qui en résulte.*

*II.-Les professionnels de santé concernés sont informés du recours à ce traitement de données. Les données du patient utilisées dans ce traitement et les résultats qui en sont issus leur sont accessibles.*

*III.-Les concepteurs d'un traitement algorithmique mentionné au I s'assurent de l'applicabilité de son fonctionnement pour les utilisateurs.»*

La norma francese 2021/1017 fa riferimento all'uso di dispositivi medici comprendenti uno specifico sistema di intelligenza artificiale integrando, nella parte quarta del codice della sanità pubblica francese (dedicato alle «Professioni sanitarie»), articolo L. 4001-3 che regola l'utilizzo da parte di un professionista sanitario nel contesto di «*un atto di prevenzione, diagnosi o cura*» di un dispositivo medico comprendente l'elaborazione algoritmica di dati appresi da dati massicci, assicurando che l'interessato ne sia stato informato e che sia, nel caso, avvertito dell'interpretazione risultante.

L'articolo 17 della Loi de Bioéthique 2021/1017 diventa la nostra base di partenza per interrogarci sulle criticità dei sistemi di intelligenza artificiale nel consenso informato.

Il primo elemento di riflessione è la sostenibilità di un consenso informato che ponga sul medico la responsabilità di informare il paziente circa «l'interpretazione risultante». Ciò conduce a due iter consequenziali. In primis, la formazione degli stessi operatori sanitari sull'uso dei sistemi di intelligenza

---

<sup>22</sup> L'art. 5 della Convenzione di Oviedo, poi, prosegue: «*Questa persona riceve innanzitutto una informazione adeguata sullo scopo e sulla natura dell'intervento e sulle sue conseguenze e i suoi rischi. La persona interessata può, in qualsiasi momento, liberamente ritirare il proprio consenso*».

<sup>23</sup> Art.1, Legge 22 dicembre 2017, n. 219.



artificiale.<sup>24</sup> Se pensiamo alla ricca casistica giurisprudenziale in tema di responsabilità medica che spesso ha giudicato su inadempimenti degli obblighi informativi<sup>25</sup>, ci rendiamo conto di come la comunicazione medico-paziente sia complicata. Ciò perché due soggetti, con due conoscenze differenti, devono interagire: il sanitario non deve solo spiegare le pratiche conformandosi alla biomedicina ma anche appurarsi che il paziente le comprenda e quindi «adeguare la comunicazione alla capacità di comprensione della persona assistita»<sup>26</sup>; ma come possiamo pretendere che un medico spieghi in modo comprensibile ed esaustivo ciò di cui non sa?

Al quesito sembra rispondere il punto III della stessa legge in cui si prevede che «i progettisti dell'elaborazione algoritmica menzionata in I assicurano che il suo funzionamento possa essere spiegato agli utenti», e cioè ai professionisti sanitari.

Ma la norma pare accentuare la criticità, perché crea una rete comunicativa fatta non più di due soggetti, professionista sanitario e paziente, bensì tre: il progettista, il professionista sanitario e il paziente. Con tre conoscenze diverse ed una sola informazione da veicolare.

A risolvere la questione è, per il momento, la teoria della garanzia di spiegabilità, cioè la capacità di fornire una ragione comprensibile all'utente umano del risultato prodotto dal sistema di intelligenza artificiale. Pertanto, se un'analisi di immagine identifica una lesione, è necessario che il medico sia in grado di spiegare se questa lesione debba essere ignorata o presa in cura.<sup>27</sup>

Nel nostro argomento di trattazione, la teoria della garanzia di spiegabilità è la capacità di dare un senso clinico ai risultati computazionali. La teoria armonizza la spiegazione del professionista sanitario e il grado di tecnicità informatica del sistema di intelligenza artificiale.

Ad esempio, per indicare se una lesione rilevata è un errore o meno, è necessario da un lato avere le conoscenze tecniche mediche e, d'altra parte, capire cosa, nel funzionamento del sistema, lo ha portato al risultato prodotto.

---

<sup>24</sup> In merito si veda C. CASONATO, *Intelligenza artificiale e medicina: l'impatto sulla relazione di cura (cenni)*, cit., p. 113. «Per dare un contenuto concretamente garantista al principio dello *Human in the Loop*, infatti, non basta prevedere genericamente la presenza dell'umano all'interno del processo di formazione della decisione medica, ma è necessario che tale figura abbia, da un lato, le competenze informatiche di base per interpretare la decisione algoritmica e, dall'altro, la volontà di svolgere un ruolo di sostanza nel percorso di diagnosi e cura».

<sup>25</sup> Come risulta da note sentenze della Corte di Cassazione, alcune qui riportate: Cass. civile, sez. III, 11/11/2019, n. 28985; Cass. civile, sez. VI, 09/12/2021, n. 39084; Cass. civile, sez. III, 07/10/2021, n. 27268; Cass. civile, sez. III, 06/10/2021, n. 27112; Cass. civile, sez. III, 25/06/2021, n. 18283; Cass. civile, sez. III, 23/03/2021, n. 8163; Cass. civile, sez. III, 16/03/2021, n. 7385; Cass. civile, sez. VI, 04/03/2021, n. 5875; Cass. civile, sez. III, 15/01/2021, n. 653; Cass. civile, sez. III, 04/11/2020, n. 24471.

<sup>26</sup> Art. 33, Codice di deontologia medica italiana.

<sup>27</sup> COMITÉ CONSULTATIF NATIONAL D'ÉTHIQUE, COMITÉ NATIONAL PILOTE D'ÉTHIQUE DU NUMÉRIQUE, *AVIS COMMUN AVIS 141 CCNE/AVIS 4 CNPEN. Diagnostic Médical et Intelligence Artificielle: Enjeux Ethiques*, Novembre 2023, in <https://www.ccne-ethique.fr/fr/publications/avis-141-du-ccne-et-4-du-cnpn-diagnostic-medical-et-intelligence-artificielle-enjeux> (ultima consultazione 5/10/2023).

Nel progetto AGIR la spiegabilità potrebbe concretizzarsi con linee guida medico informatiche o, addirittura, introdurre nuove professioni competenti nel futuro mercato del lavoro. Al fine di armonizzare la complessa interazione medicina-informatica.

Un ulteriore ordine di riflessione vive negli stessi limiti della teoria di garanzia di spiegabilità: l'effetto Black Box. Sul punto, l'attuale proposta di regolamento europeo (par. 2.3) nei «*requirements for high-risk AI system*» prevede al suo articolo 13 «*Transparency and provision of information to users: High-risk AI systems shall be designed and developed in such a way to ensure that their operation is sufficiently transparent to enable users to interpret the system's output and use it appropriately. An appropriate type and degree of transparency shall be ensured, with a view to achieving compliance with the relevant obligations of the user and of the provider*».

La disposizione è rilevante perché il progetto AGIR, in quanto SIADM<sup>28</sup>, rientra nei sistemi ad alto rischio e trova, quindi, applicazione l'art. 13 dell'AI Act che dispone per questi sistemi di un «grado di trasparenza adeguato».

L'effetto Black Box è un ostacolo ai suddetti principi, perché si manifesta nella difficoltà da parte degli utenti umani di ricostruire passo per passo il ragionamento seguito dall' algoritmo per giungere a formulare una certa previsione. È un impedimento che è accentuato dalla complessità dei recenti modelli e dipende da fattori quali il numero di reti neurali generalmente troppo elevate perché un umano riesca ad averne una visione d'insieme, a cui si aggiunge il design sub-simbolico che impedisce di fornire un'interpretazione semantica alle connessioni svolte per darne un significato comprensibile agli utenti e il contesto in cui agisce il sistema.

I nuovi sistemi di apprendimento automatico potrebbero avere una «influenza nefasta sul modo di condurre la scienza»<sup>29</sup>. Le conseguenze di una scienza senza spiegazioni ricadono - inevitabilmente - nel consenso informato perché ostacolano tutti i principi qui descritti: la comprensibile comunicazione delle informazioni, l'interpretazione del risultato e la trasparenza del sistema.<sup>30</sup>

Attualmente il problema rimane aperto ma la convivenza di sistemi di intelligenza artificiale in medicina e consenso informato potrebbe avere come ottima base di partenza il progetto AGIR. Infatti, il suo proposito di trovare un giusto equilibrio di *engagement* tra personale sanitario e sistema di intelligenza artificiale

---

<sup>28</sup> Sistema di intelligenza artificiale per diagnosi medica.

<sup>29</sup> THE ROYAL SOCIETY, ALAN TURING INSTITUTE, *The AI revolution in scientific research*, 2019, pp. 1-10, in <https://royalsociety.org/-/media/policy/projects/ai-and-society/AI-revolution-in-science.pdf> (ultima consultazione 5/10/2023). A questo proposito si veda anche: M. STEWART, *The Machine Learning crisis in scientific research*, 18 novembre 2019, in <https://towardsdatascience.com/the-machine-learning-crisis-in-scientific-research-91e61691ae76> (ultima consultazione 5/10/2023).

<sup>30</sup> C. CASONATO, *Intelligenza artificiale e medicina: l'impatto sulla relazione di cura (cenni)*, cit., p. 112.

contempla anche studiare le forme, le capacità di comunicazione umano-macchina e i gradi di apprendimento. Questi sono elementi utili a colmare un deficit che, per il momento, permette all'effetto Black Box di insediarsi tra professionista sanitario e macchina, vincolando anche il diritto del paziente di prendere decisioni libere e consapevoli.

### 3. Vantaggi e Svantaggi

I sistemi di intelligenza artificiale vengono considerati uno straordinario strumento per accelerare il raggiungimento di obiettivi di sviluppo in diversi ambiti, come quello medico, garantendo una cura più veloce e accurata.

D'altro canto, tali sistemi introducono nuove forme di esclusione e marginalizzazione, è una criticità evidenziata anche dall'UNESCO in un recente rapporto sull'Uguaglianza di Genere ed Intelligenza artificiale<sup>31</sup>. Il prossimo paragrafo analizzerà i profili di vantaggio e svantaggio.

#### 3.1. Pro: accesso alla salute e riduzione margine di errore

«La Repubblica tutela la salute come fondamentale diritto dell'individuo e interesse della collettività, e garantisce cure gratuite agli indigenti»<sup>32</sup>.

Ciascun individuo ha diritto alla salute, intesa come stato di completo benessere fisico, mentale e sociale. Il Servizio sanitario nazionale (SSN) è un sistema di strutture e servizi che hanno lo scopo di garantire a tutti i cittadini, in condizioni di eguaglianza, l'accesso universale all'erogazione equa delle prestazioni sanitarie.

L'impiego dell'intelligenza artificiale renderà più semplice ed efficace l'accesso al servizio sanitario; in particolare l'uso dell'intelligenza artificiale in quelli che sono i Livelli Essenziali di Assistenza (LEA) comporterà maggior efficienza e uguaglianza nell'erogazione del servizio in termini sia di disponibilità e utilizzo delle risorse che nelle tempistiche nell'erogazione delle prestazioni, le quali potranno essere sensibilmente ridotte. Si darà quindi piena attuazione all'art. 32 della nostra Costituzione.

Grazie all'uso di sistemi di intelligenza artificiale sarà possibile incrementare l'efficienza nel campo medico. L'introduzione porterà ad una riduzione del margine di errore poiché l'intelligenza artificiale viene definita «*noise free*»<sup>33</sup> cioè non è condizionata dalle emozioni, dalle relazioni o da interessi particolari. Questo le permetterà di svolgere un'analisi del paziente più accurata e attenta rispetto a quella che può essere fatta

---

<sup>31</sup> A questo proposito si veda UNESCO, *Recommendations on the Ethics of Artificial Intelligence*, 2022, in <https://unesdoc.unesco.org/ark:/48223/pf0000381137> (ultima consultazione 5/10/2023).

<sup>32</sup> Art. 32, c. 1, Cost.

<sup>33</sup> C. CASONATO, *Intelligenza artificiale e giustizia: potenzialità e rischi*, cit., pp. 3369-3389.

da un professionista sanitario che può avere urgenza o presentare stanchezza. Con la presenza di queste variabili il rischio di commettere un errore aumenta.

Inoltre, la macchina compiendo il lavoro in autonomia ridurrà i tempi di attesa dei pazienti perché potrà lavorare all'analisi delle immagini senza interruzione quindi il processo sarà più veloce e i pazienti potranno ottenere la loro diagnosi dopo un periodo di tempo ragionevole.

### **3.2. Contro: pregiudizio algoritmico, taglio delle code e conservative profiling**

Nella ricerca medica un utilizzo di dati, che prende atto delle differenze di genere, rende una maggiore appropriatezza della terapia e un'effettiva tutela della salute. Senza tale orientamento, la politica della salute risulterebbe discriminatoria. Ciò si amplifica nel settore sanitario che utilizza sistemi di intelligenza artificiale, il fenomeno è conosciuto come «pregiudizio algoritmico».

Per risolvere la questione dobbiamo assumere che il sistema è frutto di un'elaborazione umana ed in quanto tale produce risposte in base agli input che vengono inseriti dai programmatori. Se questi sono dati di ricerche mediche basate solo sul prototipo, ad esempio, di paziente di sesso maschile, l'algoritmo imita il nostro pregiudizio, producendo tecnologie inaffidabili perché poco rispondenti alle necessità di ampie fette della popolazione mondiale.

Uno studio del Berkeley Haas Center for Equity, Gender and Leadership ha monitorato circa 133 sistemi di intelligenza artificiale in tutti i settori dal 1988 ad oggi, scoprendo che il 44,2% (59 sistemi) dimostra pregiudizio di genere, il 25,7% (34 sistemi) mostra sia pregiudizi di genere che razziali<sup>34</sup>.

Altro grande rischio dell'uso dell'intelligenza artificiale risiede nel suo lavorare su base statistico-probabilistica utilizzando dati già forniti alla macchina.

Questo porta a due gravi implicazioni nell'ambito medico: la mancata individuazione di malattie rare e il blocco della ricerca su possibili nuove malattie.

Il primo fenomeno è noto come il «taglio delle code»: i sistemi di intelligenza artificiale tendono ad ignorare i parametri che appaiono di meno e a focalizzarsi principalmente sui parametri che appaiono con più frequenza<sup>35</sup>. La macchina, quindi, tenderà ad avere una forte propensione a basare le proprie decisioni sulle teste (rappresentate appunto da tutti quei casi sui quali il sistema ha un numero maggiore di dati) e più in

---

<sup>34</sup> BERKELEY HAAS CENTER FOR EQUITY, *Mitigating Bias in Artificial Intelligence. An Equity Fluent Leadership Playbook*, luglio 2020, in [https://haas.berkeley.edu/wp-content/uploads/UCB\\_Playbook\\_R10\\_V2\\_spreads2.pdf](https://haas.berkeley.edu/wp-content/uploads/UCB_Playbook_R10_V2_spreads2.pdf) (ultima consultazione 5/10/2023).

<sup>35</sup> Y. J. PARK, A. TUZHILIN, *The Long Tail of Recommender Systems and How to Leverage It*, in P. PAU, D. BRIDGE, B. MOBASHER, F. RICCI (a cura di), *RecSys '08: Proceedings of the 2008 ACM conference on Recommender systems*, New York, 2008, 11-18.

generale ad avere un basso livello di accuratezza per tutti quei casi rientranti nelle code (all'interno delle quali vi sono tutte quelle possibilità rappresentate da una quantità minore di dati)<sup>36</sup>.

Questo si rivela un punto critico nel momento in cui la macchina opererebbe in autonomia. Considerando poi il problema delle malattie rare, vi è il rischio che esse vengano ignorate, o, in ogni caso, non volutamente trattate dalla macchina. Naturalmente, così facendo si creerebbe una sorta di circolo vizioso: il non-trattamento di queste malattie avrebbe un impatto anche sui dati che la macchina va ad analizzare, abbassando ulteriormente la mole di dati riguardanti uno specifico caso in relazione ad altre casistiche analoghe. Questo circolo potrebbe infine concludersi con la scomparsa di fatto delle malattie che colpiscono piccole fasce della popolazione. Una soluzione data dagli ingegneri del DISI, come citato sopra, risiede nell'impiego di intelligenze artificiali che mostrano chiaramente la percentuale di accuratezza della loro risposta, lasciando alla discrezionalità del medico.

L'impossibilità dell'evolversi della ricerca su nuove patologie è legata a un fenomeno tipico dell'intelligenza artificiale: il «*conservative profiling*». È, infatti, complesso addestrare i sistemi d'intelligenza artificiale al fine di trovare delle soluzioni creative al problema che viene loro posto, piuttosto che far loro semplicemente memorizzare, e poi riproporre, una soluzione già precedentemente creata (quindi anche tenendo in considerazione nuovi parametri con i quali la macchina non si è mai interfacciata). Inoltre, anche se tale processo riuscisse, sarebbe rilevante notare che il sistema fa sempre riferimento ai parametri in precedenza forniti o assimilati, tendendo ad ignorare quelli che le si presentano *ex novo*<sup>37</sup>.

L'aspetto diventa particolarmente nocivo nel momento in cui i dati si riferiscono a malattie: la macchina non riuscendo ad uscire da quelli che sono i parametri già impostati, non riuscirebbe ad identificare (e quindi trattare correttamente) eventuali nuove malattie. Il tema è particolarmente critico per quanto riguarda la ricerca: se infatti ci si basasse unicamente sul sistema di intelligenza artificiale, lo sviluppo in campo medico rischierebbe di arrestarsi, proprio perché si lavorerebbe unicamente su dati già acquisiti in precedenza, ricondotti ad output già forniti in precedenza.

Il DISI ne ha trovato una soluzione tramite l'utilizzo di sistemi che «*sanno di non sapere*». Nel momento in cui quindi vi sono dei sintomi che presi complessivamente non porterebbero ad un output che la macchina conosce esattamente, essa procede ad arrestarsi e dà il controllo totale al medico.

---

<sup>36</sup> Y. ZHANG ET AL., *Deep Long-Tailed Learning: A Survey*, in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15 aprile 2023, 1-20, in <https://arxiv.org/pdf/2110.04596.pdf> (ultima consultazione 5/10/2023).

<sup>37</sup> M. MARCINOWSKI, *Artificial Intelligence or the Ultimate Tool for Conservatism*, in *DANUBE*, 1, 2022, pp. 1-12.

#### 4. Sostenibilità del progetto e prospettive future

Possiamo, innanzitutto, affermare che il progetto trovi riscontro e sia compatibile con molteplici discipline nonostante la regolamentazione sia recente ed ancora in elaborazione.

Inoltre, alla luce di quanto trattato, possiamo individuare due ordini di problemi maggiormente rilevanti e tentare di elaborare delle proposte migliorative.

Il primo riscontro ruota intorno alla costante necessità dell'intervento umano nei meccanismi dei sistemi di intelligenza artificiale. L'uomo ha, infatti, la funzione di monitorare il lavoro svolto dalla macchina e di confermare la decisione presa. L'obiettivo del progetto AGIR - trovare un equilibrio tra medico e sistema di intelligenza artificiale - potrà contribuire in modo significativo alla regolazione stessa dei sistemi.

Il secondo punto su cui è doveroso argomentare guarda alla trasparenza del sistema: il principio, così come delineato nell'AI Act, risulta poco rappresentativo del contesto di applicazione e di nascita dei sistemi. È un obbligo di trasparenza che pare non disponga di effettive garanzie alla, già citata, teoria di spiegabilità (par. 2.4): l'articolo 13 dell'AI Act prevede meri obblighi informativi in capo al fornitore al fine di consentire agli utenti - e cioè i professionisti sanitari, nel caso di specie - di interpretare l'output del sistema e utilizzarlo in modo adeguato.

L'obbligo di trasparenza, così come formulato, pare non farsi pienamente carico del problema Black Box. Alla luce di quanto elaborato ed in modo conforme alla stessa teoria di spiegabilità, non sarebbe maggiormente idoneo disporre di obblighi collaborativi-informativi, piuttosto che meramente informativi, tra fornitori e utenti?

La tecnica legislativa attuabile ai sistemi di intelligenza artificiale, non nella sola pratica sanitaria ma anche ai restanti plurimi settori di applicazione, potrebbe armonizzare il principio con delle linee guida accessorie, con la funzione di delineare un quadro legale altamente tecnico in cui inserire il sistema di intelligenza artificiale trasparente. La struttura qui proposta seguirebbe il generale iter giuridico procedurale, così come delineato anche nell'AI Act, che posiziona le linee guida di trasparenza in allegati con delega di competenza circa la formulazione ad un organo tecnico. Pertanto, queste saranno facilmente revisionabili e capaci di seguire, o addirittura anticipare, l'innovazione dei sistemi di intelligenza artificiale.

Concludendo, è giusto riflettere su prospettive future. Nella disciplina attuale non abbiamo spazio per un sistema di intelligenza artificiale operante in completa autonomia, ma questo costituisce un ostacolo prettamente giuridico: non è, infatti, così lontano dall'immaginario comune pensare a una macchina che guidi o operi autonomamente. E forse non dovrebbe essere altrettanto lontano riflettere su collaborazioni multidisciplinari come la medicina, l'ingegneria e la giurisprudenza per anticipare, e non rincorrere, lo sviluppo tecnologico e scientifico.

*De iure condendo*, ravvisiamo che sia necessario legiferare sui possibili autonomi spazi operativi dei sistemi di intelligenza artificiale, regolando l'applicazione automatizzata piuttosto che delegittimandola. E in realtà, *de iure condito*, già qualcosa si intravede, infatti: lo stesso art. 22, par. 1 del GDPR nella parte in cui dispone che l'interessato ha il diritto di non essere sottoposto a una decisione basata unicamente sul trattamento automatizzato vede anche al paragrafo successivo i tre casi in cui, invece, il sistema di intelligenza artificiale può operare in modo automatizzato<sup>38</sup>.

---

<sup>38</sup> L'intero articolo 22, Reg. (UE) 2016/679 qui riportato:

«1. L'interessato ha il diritto di non essere sottoposto a una decisione basata unicamente sul trattamento automatizzato, compresa la profilazione, che produca effetti giuridici che lo riguardano o che incida in modo analogo significativamente sulla sua persona.

2. Il paragrafo 1 non si applica nel caso in cui la decisione:

- a) sia necessaria per la conclusione o l'esecuzione di un contratto tra l'interessato e un titolare del trattamento;
- b) sia autorizzata dal diritto dell'Unione o dello Stato membro cui è soggetto il titolare del trattamento, che precisa altresì misure adeguate a tutela dei diritti, delle libertà e dei legittimi interessi dell'interessato;
- c) si basi sul consenso esplicito dell'interessato.

3. Nei casi di cui al paragrafo 2, lettere a) e c), il titolare del trattamento attua misure appropriate per tutelare i diritti, le libertà e i legittimi interessi dell'interessato, almeno il diritto di ottenere l'intervento umano da parte del titolare del trattamento, di esprimere la propria opinione e di contestare la decisione.

4. Le decisioni di cui al paragrafo 2 non si basano sulle categorie particolari di dati personali di cui all'articolo 9, paragrafo 1, a meno che non sia d'applicazione l'articolo 9, paragrafo 2, lettere a) o g), e non siano in vigore misure adeguate a tutela dei diritti, delle libertà e dei legittimi interessi dell'interessato».