



UNIVERSITÀ
POLITECNICA
DELLE MARCHE

FACOLTÀ DI MEDICINA E CHIRURGIA

MASTER DI I LIVELLO IN BIOETICA PER LE SPERIMENTAZIONI CLINICHE E I COMITATI ETICI

Bioetica dell'Intelligenza Artificiale e sperimentazioni cliniche delle Terapie Digitali

Relatore:

Prof. Maurizio Battino

Tesi di:

Dr.ssa Ilaria Rossiello

DATA 05/07/2024

INDICE

1) Rivoluzione digitale	pag.3
2) Definizione di intelligenza artificiale	pag.14
3) Principi etici dell'Intelligenza Artificiale	pag.18
4) I rischi etici	pag.21
5) I reali pericoli	pag.24
6) Ambiti applicativi in medicina	pag.28
7) Terapie digitali	pag.31
8) Conclusioni	pag.34
9) Bibliografia	pag.35
10) Ringraziamenti	pag.37

1) Rivoluzione digitale

Durante gli ultimi decenni abbiamo assistito a una rivoluzione epocale nella storia dell'Umanità, comparabile all'invenzione della ruota e come essa irripetibile: il passaggio dall'analogico al digitale.

La mia generazione è stata investita da questo passaggio, pertanto i ricordi della mia infanzia comprendono messaggi lasciati in carta e penna sul tavolo della cucina mentre ad oggi mi risulta impossibile concepire di non poter contattare mio marito per programmare la cena.

La generazione dei nativi digitali è invece nata in un Mondo in cui la tecnologia è parte integrata del vissuto quotidiano: dalla sveglia sul telefono a Siri che risponde alle domande sul meteo.

I risvolti pedagogici e sociali di questa Rivoluzione sono di portata mastodontica, ma non essendo l'argomento principe della mia tesi, mi limito ad accennare alle altre dimensioni che la tecnologia domina oltre a quella prettamente sanitaria.

Attualmente esistono più dispositivi digitali che persone in carne e ossa nel Mondo, essi interagiscono tra di loro anche in maniera autonoma. Potremmo immaginare il nostro Pianeta come popolato da esseri umani e macchine, ed esse rappresentano il maggior numero di contatti che quotidianamente vengono scambiati.

Un qualsiasi innocuo Smartphone, che ciascuno di noi tiene in tasca in questo momento, possiede una potenza di calcolo maggiore di quanto avesse a disposizione la NASA quando Armstrong sbarcò sulla Luna.

L'Apollo 11 aveva un processore con una RAM di 32.768bit e 72KB di ROM. Al giorno d'oggi usiamo comunemente RAM 4GB e ROM 152GB.

La mente umana non è in grado di avvicinarsi neanche lontanamente alla velocità con cui una macchina processa le informazioni che le diamo.

Un'accelerazione così repentina della velocità di processazione e dell'ampiezza di memoria ha permesso di traslarci da un Mondo gestito da hardware e bulloni a un'infosfera fatta di dati e software.

Per Infosfera intendiamo quel Mondo fatto da informazioni che viaggiano su piattaforme virtuali, prime tra tutte i social.

L'infosfera ci permette di essere in più luoghi contemporaneamente e di parlare con platee inarrivabili fisicamente. Per esempio in questo momento sto scrivendo la mia tesi a Civitanova Marche, ma tra un secondo potrei avere una chiamata con una ricercatrice olandese oppure caricare un post che verrà letto in America.

Sull'Infosfera si crea un'identità parallela alla nostra, creata proprio dalle informazioni che (più o meno consapevolmente) regaliamo alla “grande vasca”.

Su Internet si crea un nostro clone, che grazie all'intelligenza artificiale potrebbe avere anche il nostro volto e la nostra voce. Questo clone si plasmerà in base alle nostre caratteristiche individuali e alle scelte che compiamo.

Questa possibilità ha portato a immaginare scenari in cui il clone virtuale di personaggi ammirevoli (come Dante, Galileo o Steve Jobs in futuro) potranno essere interrogati per trovare delle soluzioni in un presente in cui essi non esistono più.

Altri scenari più crinosi immaginano come un clone possa estorcere denaro alle famiglie e/o creare patologie psichiatriche da non rielaborazione del lutto.

Lasciando da parte queste considerazioni, ora non si pone il problema di quali dati far emergere ma di quali dati eliminare poiché il flusso in ricezione è mastodontico e in gran parte non utilizzabile. Ogni giorno vengono caricati e processati dati in quantità inimmaginabile.

Stiamo passando da una cultura della registrazione a una cultura dell'eliminazione. Stiamo passando dalla logica alla statistica.

Le vecchie intelligenze artificiali (d'ora in poi IA) avevano il compito di produrre calcoli e compiti in modo veloce e automatizzato, le nuove IA hanno il compito di confrontare dati, imparare da essi e dare risposte che risolvano i nostri problemi.

Non ci accontentiamo più di avere dei grandi fogli di calcolo che forniscono dati grezzi e non utilizzabili, abbiamo bisogno di una tecnologia che si faccia carico di compiti che prima venivano affidati all'Uomo.

Facciamo un passo indietro per comprendere quanto l'Umanità si trovi di fronte a una nuova epoca con sfide da affrontare e pericoli da fronteggiare.

La storia dell'IA inizia nel secolo scorso e possiamo distinguere un personaggio particolarmente significativo legato alla sua nascita: il matematico Alan Turing, probabilmente noto per i film che narrano l'invenzione del macchinario "Bomba" utilizzato per decrittare i messaggi dei nazisti prodotti dalla macchina "Enigma" ("The imitation game" diretto da Morten Tyldum tratto da "Alan Turing. Storia di un enigma" di Andrew Hodges).

Si tratta di una storia vera, ma non tanto della vittoria di un congegno di IA su un altro, piuttosto della possibilità di una macchina di fornire risposte che la mente umana non è in grado di trovare. Ciò fu possibile solo grazie a degli stralci di codici trovati su navi rese prigioniere, non puramente al funzionamento di Bomba.

Nell'arco della Seconda Guerra Mondiale, Turing cambia idea sulle potenzialità dell'IA: i primi scritti (1936) descrivono il funzionamento dei calcolatori, identificando macchine in grado di svolgere compiti automatici e senza necessità di intuizione o ingegno.

Cosa intendiamo con Intuizione? La capacità di generare giudizi spontanei che non derivano direttamente da ragionamenti deduttivi.

L'Ingegno permette di scomporre passaggi e fasi che possono essere facilmente materia di programmazione informatica.

Nel 1931 Godel, tramite i Teoremi dell'Incompletezza, annunciò che qualunque sistema formale (matematico o computazionale) non sono dimostrabili all'interno del sistema stesso.

In matematica occorre affidarsi ai Dogmi, che vengono dati per Veri e non necessitano di dimostrazione. Quindi i sistemi sono autoreferenziali.

Al di là dei Dogmi, come definiamo qualcosa Vero? Quando è dimostrabile.

Ma l'Intuizione umana ci permette di capire che vero e dimostrabile iniziano a viaggiare su binari diversi, così come la mente umana e quella digitale: le mente umana ha un'intuizione che l'IA non ha, per cui vede cose che l'IA non può vedere.

Questo è uno dei motivi per cui non dobbiamo temere che l'IA crei dei cyborg che soppianteranno l'Umanità poiché dotati di un'Intelligenza umana, proprio perché mancano di Intuizione.

Facciamo un esempio: a un certo punto della Storia è arrivata l'Intuizione umana che le nostre cellule aveva un messaggio di codifica identificato poi nel DNA. Si sono iniziati a interrogare su come venisse costruito il fenotipo di un essere vivente e si è andati a cercare nel microscopico. Questa intuizione non può essere semplicemente frutto di una macchina, che lavora con altri tipi di connessioni a partite da logiche, dati e percorsi.

E' seppure vero che non abbiamo la certezza che grazie al Machine Learning (ML) e a Deep Learning (DL) prima o poi qualche fenomeno intuitivo non emergerà dalla macchina, che però ancora non avrebbe la componente emotiva e emozionale dell'Uomo.

Prima di Godel, anche Hilbert voleva dimostrare che la matematica (e quindi la programmazione) è un sistema chiuso governato dalle sue regole. Fu lo stesso Hilbert a lanciare la sfida ai matematici del Secondo Congresso Mondiale di Matematica (Parigi, 1900) di dimostrare che la matematica è completa, coerente e decidibile.

Di fatto Turing confutò che la matematica fosse decidibile inventando un computer (poiché risolve i problemi in modo autonomo).

Nei suoi scritti del 1950 (“Computing Machinery and Intelligence”) Alan Turing solleva l’ipotesi che le macchine possano accedere all’intuito se solo fossimo in grado di trovare gli enunciati che sottendono a questa abilità, implicando che la nostre menti non sono altro che macchine particolarmente complesse.

Questa è un’idea che semplifica estremamente le complesse e fini abilità della nostra specie, difatti i tentativi di insegnare alla tecnologia a essere come l’Uomo sono andati falliti.

Nel pieno della fiducia per la comprensione e la semplificazione si è cercato di scomporre l’apprendimento e quindi l’insegnamento delle caratteristiche umane, con ben poco successo.

Esistono addirittura delle teorie su come nel nostro cervello sia già insito il linguaggio, e che quindi gli idiomi esistenti non siano altro che la proiezione reale di un’abilità intrinseca nella nostra specie.

Un bambino italiano messo davanti a cartoni e babysitter inglesi nell’arco di un paio di anni diventerà perfettamente padrone della lingua senza che nessuno gli abbia mai tradotto un termine. Non accade la stessa cosa con una macchina se non impostiamo degli algoritmi di traduzione.

Bletchley introduce un concetto boicottato da Turing: la componente sociale e situazionale dell’Intelligenza.

L’Intelligenza umana è sempre definita all’interno di un sistema culturale da cui attinge ed esiste delimitata da un ambiente. La maggior parte dell’intelligenza umana risiede nella società, è esternalizzata.

Questa complessità deve essere riconosciuta prima di lavorare su sistemi di IA che siano in grado di superare il test di Turing e agire in modo coerente alla situazione in cui vengono utilizzati.

A Turing dobbiamo la nascita di questa Scienza e considerevoli scoperte, ma abbiamo superato ormai la concezione di computer come puro problem-solving algoritmico.

Un collega di Turing, Jack Good, coniò l'idea di una Ultra-intelligenza in grado di superare le qualità intellettive dell'uomo poiché in grado di migliorare esponenzialmente se stessa.

Questo scenario ha stimolato la fantasia di numerosi sceneggiatori, proponendo film apocalittici e la fine del genere umano (o perlomeno del suo predominio).

La possibilità (non da escludere) che questo salto avverrà è alla base della paura che il tema IA incute in molti, professionisti inclusi. Questo salto può essere definito come Singolarità.

La rimpiazzabilità da parte delle macchine anche degli operatori professionali viene vissuta spesso con ritrosia.

I software ad ora sono in grado di sviluppare le competenze e capacità ma non ancora di fare quel salto evolutivo di cui si teme. Neanche l'uomo è in grado di prevederlo e identificarne i tasselli necessari.

Il fatto che lo stesso Elon Mask vada a esplicitare le proprie preoccupazioni apocalittiche fa leva sulla massa e genera un sentimento diffuso di timore e avversione.

Certo è che l'IA vive una stagionalità fatta di inverni gelidi e ostili e di primavere di rinascita e ottimismo.

Di fatto, però, le IA permeano il nostro Mondo: dai pagamenti elettronici a questo stesso master online.

Chat-GPT è già sul mercato con edizioni sempre più elaborate e con dati di utilizzo da far girare la testa ai colossi industriali.

In questo studio ⁽¹⁾ sono stati confrontati studenti di medicina con la vecchia versione di GPT, portando a casa ancora una volta una vittoria da parte dell'uomo ma solo di quelli più avanti nel percorso di studi.

E' prevedibile che migliorando la GPT si vada incontro a una riduzione del margine tra i concorrenti se non a un ribaltamento del risultato.

Nel Mondo del gioco l'Uomo è stato battuto già da tempo (Deep Blue, il computer progettato dall'IBM batté il campione di scacchi Garry Kasparov già nel 1996).

L'attuale sfida tra uomo e IA sembra essersi spostata sul campo del dibattito: l'IBM ha organizzato un dibattito nel 2019 tra l'IA Miss Debater e il campione mondiale di dibattiti Harish Natarajan.

Nonostante la vittoria attribuita dalla giuria al professore, l'IA ha tenuto 25 minuti di dibattito incluse delle note umoristiche.

L'IA sta quindi prendendo piede anche per quanto riguarda la creatività e l'eloquio emotivo.

Riguardo alla creatività, chat-GPT è già in grado di comporre poesie, saggi e canzoni (sollevando problemi di paternità e copy-writer).

Eccone una dimostrazione... Ho appena chiesto a chat-GPT di scrivere la mia introduzione delle tesi sull'IA in 5 righe ed ecco la risposta:

“L'intelligenza artificiale (IA) rappresenta una delle più rivoluzionarie innovazioni nel campo della tecnologia. Con la capacità di apprendere, adattarsi e prendere decisioni autonome, l'IA sta avendo un impatto significativo in diversi settori, come l'industria, la sanità e l'automazione. Tuttavia, l'integrazione e l'adozione dell'IA comportano anche sfide e questioni etiche, che richiedono una riflessione approfondita. Questa tesi si propone di esplorare il ruolo e le implicazioni dell'intelligenza artificiale nella società contemporanea e analizzare le conseguenze positive e negative che essa porta con sé.” In mezzo secondo è stata in grado di scrivere quello che io ho prodotto dopo ore di lettura e di scrittura.

La rimpiazzabilità nel Mondo del lavoro da parte delle macchine è forse il più sentito tema legato alle IA, ma sta iniziando a entrare il fantasma della rimpiazzabilità anche nel mondo della Formazione.

Sarebbe anacronistico pensare che possiamo continuare a formare e valutare gli studenti allo stesso modo: qualsiasi compito tradizionale può essere svolto dall'IA in pochi minuti e forse anche più correttamente rispetto all'alunno. In alcuni casi è palese l'intervento della macchina, come

racconta su TikTok un'insegnante inglese di un compito affidato a ragazzi molto giovani e acerbi nell'uso della lingua in cui al classico compito di writing il bambino ha proposto una lettera che suona più meno così: "mi duole informarla che non potrà essere presente al nostro scambio culturale" rivolta a un bambino di 10 anni...

Non è detto però che ragazzi più grandi non usino l'IA per scrivere temi, tradurre saggi e risolvere compiti matematici.

Un altro tipo di preoccupazione nell'ambito lavorativo è la sicurezza nel posto di lavoro. Pochi mesi fa un operaio in Oriente è stato ucciso da una macchina che lo ha scambiato per una lattina di pomodori.

La disoccupazione tecnologica è realmente un rischio o stiamo vivendo un passaggio simile a quello della Rivoluzione Industriale, ovvero la sostituzione dell'Uomo in alcuni compiti faticosi e non gratificanti a favore di una maggiore produttività e libertà di espressione delle capacità umane? C'è una grossa differenza tra quell'Epoca e quella attuale: mentre in Passato si è passati da una leva meccanica che sollevasse l'Uomo da un lavoro fisico, l'IA è una leva cognitiva che vuole sostituire l'Uomo nelle sue funzioni cerebrali superiori.

L'IA non è però una novità: ufficialmente la disciplina nacque nel 1956 durante la conferenza di Dartmouth, nel New Hampshire, dove un gruppo di matematici si riunì per parlare dell'Intelligenza Artificiale (nome coniato giusto l'anno prima da John McCarthy, professore di matematica proprio al Dartmouth College).

McCarthy portò avanti un progetto che aveva l'ambizione di descrivere l'apprendimento e l'Intelligenza Umana in modo così dettagliato da poter essere trasferito all'interno di una macchina.

L'enfasi fu grande ma andò a scemare negli anni '90 a causa di una non raggiungibilità dei risultati che si erano prefissati.

Per decenni infatti l'IA è stata più materia di studio e ricerca che reale materia applicativa.

Una nuova fulgente epoca è iniziata con l'avvento del World Wide Web e dei Big Data.

Improvvisamente all'AI iniziano a confluire dati su dati, interazioni su interazioni su scala mondiale.

Il test di Turing è stato infine superato: la IA ricorre a stratagemmi per coprire le sue aree di non funzionamento attingendo a schemi di comportamento umani.

Il web diventa una piattaforma florida di studio e ricerca applicata all'IA, con particolare interesse verso i risvolti commerciali e il linguaggio naturale.

Il termine "Big Data" fa la sua comparsa nel 1997 in un documento della NASA, ma inizia a far discutere nel 2014 con l'articolo di Forbes "12 big data definition: what's yours?".

Addirittura l'amministrazione Obama si è dotata di una figura responsabile della gestione dei big data.

GAFAM (Google, Apple, Facebook, Amazon e Microsoft) e BATX (Baidu, Alibaba, Tencent e Xiami) sono rispettivamente i draghi americani e cinesi. Sono loro a dominare le tecnologie digitali e le "grandi vasche" di dati. Questi giganti hanno in mano i Big Data e sono stati solo parzialmente regolamentati in tema di privacy dal GDPR.

E' quindi solo di recente che l'IA è passata da essere materia di ricerca a una realtà che ha stravolto prepotentemente la vita del singolo e della collettività.

Nei primi stadi di IA ci si rifà a un'idea di homo faber piuttosto che homo sapiens, quindi un'idea ottocentesca di uomo costruttore (Compte) piuttosto che privilegiare la componente arendtiana di saggezza e conoscenza.

La deviazione verso la tecnica può far perdere forza alla saggezza e all'epistemologia, che è un retaggio della Rivoluzione scientifica.

La possibilità di usare il Deep Learning e l'idea che emerga una Singolarità porta ciclicamente i riflettori sull'IA e sulle sue potenzialità costruttive e distruttive.

L'IA può essere considerata una tecnologia di terzo ordine: si situa tra noi e la natura (primo ordine) ma anche tra noi e un'altra tecnologia (secondo ordine).

Un'altra peculiarità dell'IA è che non si limita a frammentare o potenziare la realtà, ma la trasforma.

Stiamo costruendo una Società che si adatta alla tecnologia in uso: le forme delle case e dei giardini agevolano il lavoro dei robot dediti alla pulizia. Le cucine si adattano agli strumenti in uso: non c'è un robot dalle sembianze umane che prende i piatti dal lavabo, li lava e li rimette a posto ma ci sono cucine costruite per ospitare una lavastoviglie che inglobi i piatti e esegua il compito assegnato. Alla fine i piatti sono stati lavati, ma modellando l'ambiente in funzione della macchina.

Questo processo si estende anche alle nano e biotecnologie, che stanno re-ontologizzando e re-epistemologizzando il Mondo, cosa che le tecnologie di primo e secondo ordine non sono in grado di fare.

La nostra Vita non è più solo qui e ora, ma è potenzialmente ovunque e in qualsiasi tempo: possiamo essere fisicamente nelle Marche ma parlare in streaming a una platea di medici a New York. Possiamo programmare a che ora far partire l'irrigazione del prato o pubblicare un post.

Allo stesso modo la tecnologia ha creato uno scollamento della Realtà su vari livelli. A partire dall'Identità: prima dell'analogico si avevano documenti cartacei posseduti fisicamente dal soggetto, ora abbiamo documenti e password virtuali come anche la profilazione delle informazioni che fanno capo al soggetto. Non abbiamo solo l'Identità su Internet, ma una mole di dati personali forniti più o meno coscientemente dal soggetto stesso.

Ci sono app in grado di creare la nostra immagine corporea tra decenni (famoso in trend di Facebook di qualche anno fa di invecchiamento

corporeo, che altro non era che un fenomeno creato ad hoc per raccogliere immagini umane reali).

Ci sono app che sono in grado di far parlare i morti (ovvero creare un video dell'avatar della persona deceduta). I possibili sviluppi criminali sono solo immaginabili. Il lato più inquietante della vicenda è che siamo stati noi a regalare le immagini necessarie all'IA per acquisire queste capacità solamente aderendo a delle challenge sui social media.

Questa esplosione di complessità e il rischio di scollamento tra Identità e Soggetto ha portato anche alla regolamentazione europea GDPR per proteggere la privacy dei soggetti sul trattamento in senso lato dei dati sensibili.

Per quanto i dati vengano criptati e le password diventino sempre più complesse, anche le IA vengono allenate a una maggior capacità di decodifica, cosicché è diventato necessario aggiungere alle password l'identificazione biometrica (impronta digitale o lineamenti del viso) oppure la doppia verifica.

L'IA ha ancora bisogno dell'Uomo per decifrare semanticamente dei contesti e affinare il proprio funzionamento, per questo l'Uomo viene sempre maggiormente usato per migliorare l'IA (computazione basata sull'umano), in questo modo ci stiamo sempre maggiormente fondendo con essa e creando legami a doppia mandata.

È famosa la frase “se non stai pagando per qualcosa, il prezzo sei tu”: usando le nostre azioni e i nostri dati il sistema evolve, ecco perché la maggior parte dei servizi sul web è gratuito, stiamo pagando con le nostre azioni umane oppure tramite la cessione dei dati che verranno usati per renderci clienti e creare interfacce fluide.

2) Definizione di Intelligenza Artificiale

Si parla sovente e in diversi ambienti scientifici di Intelligenza Artificiale, ma capita di riferirsi con questo termine a un insieme indistinto e mal definibile di strumentazione e algoritmi senza avere una condivisione dell'Idea che abbiamo di questo termine.

In tal modo i due (o più) partecipanti al dibattito potrebbero parlare a lungo senza arrivare a nessuna analisi gratificante o avendo il dubbio di parlare dello stesso tema.

In realtà, la confusione e il dibattito coinvolgono anche il termine stesso di Intelligenza Artificiale.

Avendo scelto la parola "Intelligenza", ci troviamo nell'imbarazzo di definire cos'è l'Intelligenza e nell'ancor maggiore imbarazzo di ammettere che "Intelligenza Artificiale" è un ossimoro poiché le macchine non hanno ancora fatto il salto evolutivo necessario a guadagnare un'Intelligenza pari a quella umana per caratteristiche intrinseche.

L'IA scinde l'esecuzione di un compito dalla capacità di usare intelletto, pertanto un compito che può essere eseguito dall'IA si dimostra essere dissociabile dall'Intelligenza.

Non esiste una definizione ufficiale di IA, proprio perché non è un'entità finita e definita ma un insieme di discipline, prodotti tecno-scientifici e algoritmi.

Le definizioni esistenti ruotano intorno ai seguenti 4 concetti (2):

- Sistemi che pensano come gli umani;
- Sistemi che agiscono come gli umani;
- Sistemi che pensano in modo razionale;
- Sistemi che agiscono in modo razionale.

A questo punto è necessario riparlare del famoso test di Turing, geniale nella sua semplicità.

Questo test mira a capire se una macchina può avere un comportamento intelligente tramite il giudizio di un osservatore esterno sul suo operato: se non siamo in grado di riconoscere che l'output fornito è stato generato da una macchina allora il test è stato superato.

“Propongo di considerare la domanda: possono le macchine pensare? Questa indagine dovrebbe iniziare definendo il significato di termini come macchina e pensare... La domanda possono le macchine pensare credo sia troppo insensata per meritare di essere discussa. (Turing, 1950)”.

Quindi valutiamo la capacità di fare qualcosa piuttosto del suo modo di farla.

La macchina può superare il test se possiede una perfetta padronanza della lingua, un bagaglio di conoscenze, ragionamenti automatizzati e la capacità di imparare all'interno del suo ambiente.

A questo proposito, occorre fare chiarezza su due sottoinsiemi dell'IA che sono il Machine Learning (ML) e il Deep Learning (DL).

Per ML intendiamo la branca di IA che si pone come obiettivo quello di far sì che i computer migliorino la loro performance basandosi sull'esperienza. Quindi alla macchina non viene detto come rispondere a una determinata domanda, ma le viene insegnato a dare una risposta e a perfezionarsi nel processo.

Vengono volutamente introdotti dei bias per istruire il sistema a correggerli. Esistono diversi metodi di apprendimento: con o senza supervisione, semi-supervisionati o mediante premi di rinforzo.

Sembra fantascienza, ma è su questo che ogni giorno nella nostra casella mail vengono bloccate e messe in spam decine di mail grazie a un processo statistico di selezione delle parole spammy e dei format spam, che vengono costantemente messi alla prova dai provider.

Una definizione più dettagliata di ML venne fornita nel 1997 da Mellon Tom Mitchell: “si dice che un programma informatico impara dall’esperienza E rispetto a una classe di task T e alla misurazione della performance P, se la sua performance nei task T come misurato da P, migliora con l’esperienza E”.

La sua natura simulativa non permette il progresso verso l’Intelligenza generale ma solo applicazioni specifiche. Quindi dipende dalla nostra capacità di semplificare e riprodurre dei processi al computer.

Con il DL invece andiamo a simulare il cervello umano mediante una rete neurale artificiale. Può essere considerato un approccio di apprendimento supervisionato.

Questa rete apprende dati e cresce in base alle sue capacità computazionali in maniera autonoma, con la seria problematica di perdita di trasparenza del processo e di distorsione in base ai dati inseriti.

Di questo ci occuperemo nel capitolo dedicato ai potenziali rischi associati all’IA. Un esempio di DL è quello dei traduttori simultanei.

L’altra faccia della medaglia è rappresentata dai Big Data: è solo tramite l’unione della potenza algoritmica e la mole di informazioni che è possibile sfruttare al massimo le potenzialità di ciascuno dei due componenti. L’uno senza l’altro non sarebbe in grado di performare ad alto livello.

La base su cui si fonda il lavoro dell’IA è l’algoritmo, che possiamo definire come una sequenza finita e ordinata di operazioni non ambigue ed effettivamente computabili che, quando viene eseguita per risolvere un problema specifico, riceve un dato in ingresso (input), esegue delle elaborazioni, si arresta in un tempo finito e produce un dato come risultato (output).

In ultimo, riportiamo la definizione del Comitato Nazionale per la Bioetica 2020: “Non è scontato dare una definizione omogenea di Intelligenza Artificiale (IA) soprattutto alla luce dei recenti sviluppi che hanno portato ad un uso molto diffuso del termine. In generale, con tale termine può identificarsi il settore delle Tecnologie dell’Informazione e della

Comunicazione (ICT, Information and Communication Technologies) che ha l'obiettivo di imitare con tecnologie informatiche alcuni aspetti dell'intelligenza umana, per sviluppare -prodotti informatici o macchine- in grado sia di interagire e di apprendere dall'ambiente esterno, sia di assumere decisioni con crescenti gradi di autonomia.”

I Chat-bot sono dei software che utilizzano il linguaggio naturale e il ML per interfacciarsi con l'utenza fornendo delle risposte alle domande. Possono essere dichiarativi e dare risposta a comuni query in forma colloquiale oppure assistenti virtuali più complessi. Utilizzano il DL per imparare dal comportamento dei precedenti utenti e quindi implementare un'Intelligenza predittiva.

Siri e Alexa sono Chat-bot che memorizzano e profilano l'utente fornendo suggerimenti personalizzati, non è un caso che Amazon ci fornisca titoli sempre allineati ai nostri gusti (con il rischio di entrare in echo-chamber oltre a quello di non aprirsi mai a nuovi stimoli).

Così come non è casuale che conversando con una nostra amica di pantaloni per la palestra ci si riempiano siti e social di pubblicità proprio su suddetto abbigliamento: il microfono è sempre in ascolto e a lavoro.

Nel 2018 l'azienda OpenAI lancia il primo GPT (Generative Pre-trained), ma nel 2020 esce sul mercato la sua terza versione che contiene 175 miliardi di parametri.

ChatGPT è in grado di creare un testo in modo fluido e attingendo le informazioni dalla rete. Questa IA è stata addestrata sulla probabilità di avere una parola in seguito a un'altra in un testo.

Ci avviciniamo all'intuizione di cyborg (cybernetic organism) di Manfred Clyones e Nathan Kline, che negli anni '60 hanno immaginato una macchina in grado di simulare se non provare addirittura emozioni e sentimenti. Nella cinematografia i cyborg appaiono sovente, come in Intelligenza Artificiale di Steven Spielberg del 2001 o nella più recente serie Westworld del 2016.

3) Principi etici dell'Intelligenza Artificiale

Siamo di fronte a un nuovo modo di agire di fronte a un compito per raggiungere un obiettivo.

Accantonando gli scenari del tutto ipotetici di evoluzione dell'IA, dobbiamo porci delle domande su come le attuali tecnologie portino con sé dei dilemmi etici nuovi e non trattati dalla Disciplina.

Nonostante la Bioetica sia una branca piuttosto recente, non è del tutto noto come i suoi principi debbano declinarsi in questo contesto proprio perché la tecnologia offre ogni giorno un cambiamento di paradigma difficile da analizzare e normare in tempi rapidi.

L'IA, a dispetto del nome, porta un nuovo modo di agire che non deve essere intelligente creando un disallineamento tra agire e Intelligenza e sfuggendo alle comuni norme morali e etiche insite nell'Uomo.

Questo genera la necessità di definire nuovi principi etici.

In molti si sono cimentati nell'identificazione di questi principi, i principali sono:

- 1) Principi di Asilomar per l'IA (future of life institute), 2017
- 2) Dichiarazione di Montreal per l'IA responsabile, 2017
- 3) Ethically aligned design: a vision for prioritizing human wellbeing with autonomous and intelligent systems, IEE 2017
- 4) Dichiarazione su intelligenza artificiale, robotica e sistemi autonomi, EGE 2018
- 5) Cinque principi generali per un codice di intelligenza artificiale (House of Lords, Artificial intelligence committee), AIUK 2017
- 6) Principi di partenariato sull'IA, 2017

Poiché l'IA è una nuova forma di agire, la Bioetica è l'area dell'etica più vicina all'Etica Digitale e si propone di riflettere ecologicamente su nuove forme di agenti, pazienti e ambienti (Floridi, 2013).

Quattro dei cinque principi di etica alla base dell'IA sono difatti già presenti in Bioetica:

- 1) **Beneficenza**. L'IA è creata e utilizzata a beneficio dell'Umanità: promuove il benessere di tutte le creature senzienti (Montreal) oppure dare priorità al benessere umano come risultato in ogni design di sistema (IEE) oppure beneficiano e dotano di maggiori poteri quante più persone possibili (Partenariato) oppure assicurare le precondizioni di base per la vita sul nostro pianeta, la continua prosperità per l'Umanità e la conservazione di un buon ambiente per le generazioni future (EGE);
- 2) **Non maleficenza**. Il timore principale che si può avere nei confronti dell'IA è che questa acquisisca una sorta di coscienza e inizi in modo consapevole a fare del male al genere umano per conquistare una completa autonomia. Questo è però inverosimile fino all'avvento della Singolarità di cui si parlava in precedenza. Timori più fondati su un'involontaria lesione riguardano la privacy, la corsa agli armamenti e l'auto-miglioramento ricorsivo;
- 3) **Autonomia**. Permettendo alla tecnologia di svolgere dei compiti in nostra vece, cediamo volontariamente parte del nostro potere decisionale. Per questo motivo, tanto maggiormente svilupperemo e demanderemo mansioni all'IA, tanto maggiore sarà la nostra perdita di Autonomia. Per cui lo sviluppo dell'IA dovrebbe promuovere l'Autonomia di tutti gli esseri umani (Montreal) oppure i sistemi autonomi non devono compromettere la libertà degli esseri umani (EGE) oppure il potere autonomo di danneggiare, distruggere o ingannare gli esseri umani non dovrebbe mai essere conferito all'IA (AIUK) oppure gli esseri umani dovrebbero scegliere come e se delegare le decisioni ai sistemi di IA, per raggiungere gli obiettivi scelti dagli esseri umani (Asilomar). Possiamo definire questo nuovo concetto di

Autonomia come meta-Autonomia o modello di decisione di delega (Floridi, 2022) ovvero il potere di decidere di decidere (e la garanzia di poterlo rivedere);

- 4) Giustizia. Lo sviluppo dell'IA dovrebbe promuovere la giustizia e cercare di eliminare tutti i tipi di discriminazione (Montreal) oppure l'IA dovrebbe contribuire alla giustizia globale e alla parità nell'accesso ai benefici (Asilomar);
- 5) Esplicabilità. Questo è l'unico principio non già presente e rappresenta la necessità di comprendere i processi decisionali dell'IA. Asilomar utilizza il termine "trasparenza", mentre AIUK utilizza "intelligibilità". Questo principio include il senso epistemologico di "intelligibilità" e il principio etico di "responsabilità", permette di legare i principi di beneficenza e non malevolenza. Comprendere il funzionamento del delegato al nostro comportamento è cruciale anche per ritenere eticamente e legalmente responsabile il delegante in caso di esiti gravi e negativi. La maggior parte delle IA attualmente si può schematizzare come una black box che riceve degli input e elabora degli output. Si chiama black box proprio perché il processo di elaborazione è imperscrutabile, da qui il serio dilemma della trasparenza.

4) I rischi etici

Chiarificati quelli che sono i principi etici alla base dell'IA, vediamo in questo capitolo quasi rischi si corrono nell'utilizzo dell'IA nel contravvenire i principi in precedenza esplicitati.

Tali rischi sono trasversali e riguardano altri ambiti applicativi.

Poiché la tecnologia è ormai parte inscindibile della Società odierna, occorre spostare il dibattito su come prevenire questi rischi o gestirli al meglio.

Il passaggio da una prima fase, più teorica, relativa al cosa, a una seconda fase, più pratica, relativa al come, per così dire, è ragionevole, lodevole, fattibile ⁽³⁾.

- 1) *Shopping etico*. Il fiorire di riflessioni e documenti su quali siano i principi etici alla base dell'IA ha generato confusione e ridondanza. Nel capitolo precedente abbiamo preso in considerazione solamente le fonti più autorevoli e condivise su scala mondiale, ma sono solamente una frammento della letteratura prodotta. Lo shopping etico prevede la scelta dei principi e definizioni che maggiormente si adattano alle esigenze dello stakeholder, che cede alle pressioni sociali sulla correttezza del loro agire ma piegando il tema ai propri bisogni. La definizione dello shopping etico digitale è: il malcostume di scegliere, adattare o rivedere (“mescolare e abbinare”) principi etici, linee guida, codici, quadri di riferimento o altri standard simili (specialmente ma non solo nell'etica dell'IA) estraendoli da una varietà di offerte disponibili, per conferire, una patina nuova ad alcuni comportamenti preesistenti (scelte, processi, strategie, ecc.) e in tal modo giustificarli a posteriori, invece di implementare o affinare nuovi comportamenti confrontandoli con standard etici pubblici;
- 2) *Blue washing etico*. Molto più conosciuto è certamente il greenwashing, il blue washing è il suo alter ego in tema di IA: il malcostume di fare affermazioni infondate o fuorvianti al riguardo (o di attuare misure superficiali a favore) dei valori etici e dei benefici di processi, prodotti,

servizi o altre soluzioni digitali al fine di apparire più etici dal punto di vista digitale di quanto non si sia effettivamente. Il settore pubblicitario è parte integrante del mercato e su esso investono le grandi aziende. Ciclicamente emergono all'attenzione pubblica scandali sulla disinformazione prodotta dal marketing, ma è necessario un costante sforzo attivo per promuovere trasparenza e informazione;

- 3) *Lobbismo etico*. Si definisce nel malcostume di sfruttare l'etica digitale per ritardare, rivedere sostituire o evitare un'ideale e necessario regolamentazione giuridica (o la sua applicazione) relativa al design, lo sviluppo e l'implementazione di processi, prodotti, servizi o altre soluzioni digitali. Il rischio è quello di vicariare con l'etica la norma giuridica. In un tema tanto recente il rischio è ancor maggiore poiché la legislazione non ha molta storia alle spalle e maturità di trattazione. Il lobbismo etico è perciò destinato a sparire nel lungo termine, ma può provocare danni nel frattempo che la giurisprudenza faccia il suo corso;
- 4) *Dumping etico*. Indica l'esportazione di pratiche di ricerca contrarie all'etica di Paesi dove vi siano cornici legali ed etiche e meccanismi di applicazione delle norme più deboli (termine coniato nel 2013 dalla Commissione europea). Abbiamo affrontato proprio in questo Master un analogo rischio legato alla ricerca e alla sperimentazione clinica di farmaci, attualmente normato proprio per evitare questo rischio in ambito sanitario. Più estesamente si può definire come il malcostume di esportare attività di ricerca riguardo a processi, prodotti, servizi o altre soluzioni digitali, in altri contesti o luoghi (per esempio, da organizzazioni europee al di fuori dell'UE) in modi che sarebbero eticamente inaccettabili nel contesto o luogo di origine, e di importare i risultati di tali attività di ricerca contrarie all'Etica;
- 5) *Elusione dell'Etica*. Si definisce come il malcostume di svolgere sempre meno lavoro etico (come adempiere ai doveri, rispettare i diritti, onorare gli impegni ecc.) in un dato contesto quanto più è percepito (erroneamente) il

ritorno di tale lavoro etico in quel contesto. È molto più frequente che questo avvenga in Paesi meno sviluppati, in cui anche i controlli e le pressioni sull'Etica sono minori. Avere chiaro chi ha la responsabilità dell'agire permette di contrastare questo fenomeno.

5) I reali pericoli

Come ogni cosa, l'IA può essere utilizzata a fin di bene ma anche a scopo criminale.

Così come Nobel quando inventò la dinamite (1867) non aveva intenzione di utilizzarla a scopi bellici, l'IA non è nata per i crimini che di seguito elencherò.

I crimini di IA non sono potenziali, dovrebbero essere questi a spaventare i cittadini e rappresentano l'attuale faccia negativa della medaglia.

Questi crimini prevedono l'utilizzo dell'IA come condizione necessaria e di fatto non esisterebbero senza l'IA.

La fattibilità di alcuni crimini è stata dimostrata: Seymour e Tully nel 2016 hanno creato un link di phishing costruito grazie alla profilazione degli utenti, questo link permetteva l'accesso a informazioni facilmente utilizzabili a scopi criminali.

Sempre nel 2016, Martínez-Miranda, McBurney e Howard hanno simulato un mercato e lo hanno manipolato grazie a ordini ingannevoli.

La ricerca sui crimini di IA è trasversale a vari ambiti: dal giuridico all'informatico passando per lo psicologico e sociale.

Attraverso il ML è possibile perdere trasparenza durante il processo poiché l'IA "evolve" basandosi sugli input forniti ma continua questo processo sul prodotto dei primi risultati. È quindi possibile che azioni e piani autonomi emergano senza il controllo dell'essere umano, queste azioni potrebbe rivelarsi criminali nonostante l'intenzione iniziale del programmatore.

Un ulteriore rischio è quello di aumentare le disuguaglianze sociali a partire da un bias di raccolta del dato: in America la profilazione delle persone di colore o che hanno frequentato la galera genera output discriminanti rispetto a un corrispettivo discendente americano.

Così come le disuguaglianze di genere della nostra Società producono dati che forniti in analisi alle macchine produrranno solo un aumento delle disuguaglianze.

Per esempio se affidassimo alla macchina la scelta tra vari curricula di un candidato per un posto dirigenziale, essa sicuramente sceglierebbe un uomo caucasico.

Imprevedibilità e autonomia possono conferire un grado maggiore di responsabilità alla macchina ma possono anche renderla meno affidabile (4). A livello giuridico per avere la responsabilità dell'atto criminale occorre avere l'Actus Reus e la Mens Rea ovvero porre in essere volontariamente l'atto e avere l'intenzione di compiere l'azione mediante l'IA o, in ogni caso, sapere che l'utilizzo dell'IA potrebbe causare l'azione criminale o anche un'omissione (criminale).

L'IA quindi non può essere protagonista di Mens Rea ma solamente di Actus Reus.

Quindi un agente artificiale può essere responsabile casualmente di un atto criminale ma solo l'uomo può esserlo moralmente (5).

Il problema cruciale è la potenzialità che hanno le ML di far emergere autonomamente atti criminali fuori dalle previsioni dei programmatori. In questo modo i programmatori diventerebbero responsabili prima facie in caso di reati gravi contro persone che devono essere in qualche modo risarcite.

Contro ogni aspettativa, le IA sono in grado di conquistare la fiducia dell'utenza (6).

Sono vecchi di decenni gli esperimenti che dimostrano come gli umani cedono informazioni estremamente personali ai sistemi artificiali.

La possibilità di dare una forma antropomorfa e la manipolazione psicologica conducono addirittura all'esecuzione di reati sessuali (7).

Ci sono persone che hanno relazioni sentimentali con degli avatar governati da algoritmi. Diverso dal caso di un americano che ha programmato un Chat-bot su un app di incontri per trovare la sua anima gemella: la sua attuale moglie ha parlato con il suo avatar per mesi prima di incontrare il programmatore e, infine, sposarlo realmente.

Torniamo ad alcuni esempi di crimini:

- *Manipolazione del mercato.* Ovvero azioni o operazioni da parte di partecipanti al mercato che tentano di influenzare artificialmente i prezzi di mercato ⁽⁸⁾, anche se queste forme di inganno emergono dal ML autonomamente poiché la ML può imparare a generare segnali ingannevoli. In questo modo vengono creati ordini falsi per manipolare il mercato, ovviamente a scopo di lucro. Questa manipolazione può coinvolgere le azioni di aziende il cui valore viene pompato in modo fittizio per poi rivenderle a un prezzo maggiore rispetto al reale mercato.

- *Traffico di sostanze illecite.* Le sostanze possono viaggiare attraverso mezzi guidati interamente da remoto (es. droni o sottomarini) e quindi difficilmente sottoponibili a controlli da parte delle Forze dell'Ordine. In caso di intercettazione della sostanza, appare comunque estremamente difficile ritracciare i mandatori e i soggetti coinvolti nello spaccio. Inoltre è possibile usare i social e il deep web per pubblicizzare in modo più o meno occulto prodotti illeciti.

- *Reati contro la persona.* In ambito di tecnologia, questi coinvolgono le molestie e la tortura. Si possono configurare questi reati tramite la creazione di social bot che diffondono messaggi di odio contro una persona. Purtroppo è cronaca di quest'estate il suicidio di un TikTokker in diretta proprio per uno scandalo su reati sessuali (peraltro non commessi) che in seguito a diffusione sui social hanno creato un danno psicologico tanto grave da portare a questo estremo gesto. Questo tipo di malware può generare problemi psicologici ingenti sulla GenZ che vive sentitamente il web e a volta fatica a dissociare la propria Identità da quella che si crea sull'Infosfera. Questa potenzialità criminosa ha portato la piattaforma TikTok –che ricordiamo essere privata e pertanto perfettamente autonoma nella gestione delle linee guida comunitarie- a creare algoritmi volti a eliminare violenza e intenti criminali dalla piattaforma.

- Reati sessuali. L'IA può avere un ruolo in questo tipo di crimini in quanto aumenta l'oggettivazione sessuale e aumenta il desiderio sessuale verso questo reato. I sex-bot permettono anche la simulazione di violenze sessuali. Alcuni sex-bot sono stati progettati proprio per simulare stupri su adulti e su minori ⁽⁹⁾, in questo stesso studio si dimostra come è comune avere fantasie di stupro. Fintanto rimane una fantasia non possiamo parlare di reato, ma c'è il serio pericolo che questo utilizzo dell'IA desensibilizzi l'uomo verso reati commessi contro altri esseri umani. Attualmente il dibattito rimane aperto poiché non è dimostrato che ci sia una correlazione diretta tra uso di IA e reati sessuali fisici.

- Furto, frode, contraffazione e sostituzione di persona. I dati che viaggiano in rete possono riguardare grandi aziende come singole persone. I tentativi di phishing sono sempre maggiormente personalizzati (spear phishing) poiché 4 volte più efficaci ⁽¹⁰⁾. La creazione di un'identità parallela su altri siti sfugge più facilmente ai controlli. In ultimo l'identità può essere utilizzata per accedere al profilo bancario dell'utente. Uno dei crimini più diffusi è collegato alla carta di credito poiché necessita di pochi dati usati da remoto.

6) Ambiti applicativi in medicina

Secondo questo studio ⁽¹¹⁾, uno degli otto settori in cui l'IA avrà maggior impatto sarà proprio quello sanitario, con previsioni che parlano di un compounded average growth rate di 27.6 miliardi entro il 2025 ⁽¹²⁾.

In Medicina Generale utilizziamo già parzialmente la tecnologia: ogni medico di Medicina Generale (MMG) deve acquistare la licenza di utilizzo di un gestionale collegato con il MEF e con il SistemaTs per permettere la dematerializzazione delle prescrizioni.

Essendo gli MMG dei liberi professionisti in convenzione, la scelta del software è libera e pertanto il carico dei suoi costi è sul lavoratore. Attualmente il monopolio del mercato è di Millewin, che non funge solamente da raccoglitore di dati dei pazienti, ma permette anche l'elaborazione e l'estrazione del dato (grazie a delle estensioni come GPG è possibile far lavorare il sistema come IA).

Inoltre permette di segnalare incongruenze, bloccare prescrizioni per motivi clinici (es. allergie), incrociare i dati con il Ministero dell'Economia (per le esenzioni di reddito) e calcolare numerosi score prodotti da una testistica validata.

L'unico ambito in cui viene utilizzata l'IA in Medicina Generale è questo e neanche al pieno delle sue possibilità (basti pensare che GPT è un'estensione a pagamento che di solito utilizzano solo i ricercatori HealthSearch).

Stando alla mole di dati che transitano nei nostri ambulatori, alla real life caratteristica dei nostri interventi (pazienti con comorbidità e polifarmacoterapie) e alla possibilità di seguire longitudinalmente i pazienti, è una grossa occasione mancata non usare maggiormente le IA per migliorare i processi amministrativi, di personalizzazione dell'intervento e

di automatizzazione della prevenzione, della cura e della gestione del cronico.

A livello ospedaliero la situazione è anche peggiore: il fascicolo sanitario elettronico è funzionante solo nelle Regioni più virtuose (es. Emilia Romagna), mentre la maggior parte dei reparti viaggia ancora con carta e penna.

Escono continuamente sul mercato strumenti diagnostici sofisticatissimi, ma quasi nessuno viene poi realmente utilizzato in ospedale (es. il Da Vinci, un robot in grado di operare chirurgicamente sotto la guida del chirurgo ma con una precisione maggiore rispetto alla mano umana, anche grazie agli ingrandimenti che è in grado di produrre sullo schermo).

Sul privato invece, iniziano a fiorire app e dispositivi di ogni tipo rivolti alla cura e al benessere personale. I tech-giants stanno sempre maggiormente attenzionando il settore sanitario come ambito di espansione, promettendo una sempre maggiore digitalizzazione dei servizi.

Basta guardare alla vicina Inghilterra, non più membro UE, ma sempre una delle Nazioni con un Sistema Sanitario Nazionale (S.S.N.) più simile al nostro rispetto a molte altre.

Già prima del COVID il National Health Service (NHS), analogo britannico del S.S.N., ha avviato una collaborazione con la Babylon Health, azienda produttrice di un'app che permette all'utenza di usare un Chat-bot di triage, di effettuare una tele-consultazione con un medico (che può sfociare in un invio sul territorio per la visita de visu) e addirittura di indirizzare il colloquio tramite l'IA.

Come? Fornendo al medico indizi sullo stato emotivo dell'utente (grazie all'analisi facciale) e sulla più probabile diagnosi a fronte delle informazioni raccolte.

Inoltre Babylon registra autonomamente la cartella clinica dell'utente che ha effettuato il colloquio.

I cittadini hanno potuto scegliere se affidarsi al classico GP operante in una practice oppure all'app.

Esistono anche dispositivi high-tech tessili che permettono di monitorare in maniera continua e non invasiva numerosi dati clinici degli indossatori (pressione arteriosa, frequenza cardiaca, glicemia, temperatura corporea, atti respiratori), essi comunicano con delle app che allertano il cliente e il medico di eventuali anomalie che sottendano situazioni emergenziali.

Oltre al più diffuso e utilizzato Apple Watch, che oltre a fungere da sveglia e telefono, permette di registrare parametri clinici (es. allertando in caso di aritmie).

Il PNRR ha permesso agli Stati di investire in digitale e in strutture fisiche per riorganizzare il Servizio Sanitario tenendo in considerazione il nuovo scenario post-pandemico. La spinta verso la ricerca e l'uso del digitale è sicuramente virtuosa sebbene non sia stata sfruttata fino in fondo in tutte le Regioni.

Durante la pandemia abbiamo assistito a un profondo cambiamento nella comunicazione medico-paziente, a un maggiore utilizzo del digitale (soprattutto in termini di teleconsultazioni e prescrizioni telematiche). Ricordiamo che proprio durante lo stato di emergenza è stata emessa l'ordinanza della Protezione Civile che ha autorizzato l'invio tramite mail delle impegnative dematerializzate, poi prorogata con il decreto 1000 proroghe e infine convertito in Legge di Stato.

Si sta dibattendo proprio in questo periodo della televisita per l'emissione dei certificati di malattia INPS.

Difatti attualmente un lavoratore malato deve necessariamente essere visitato di persona per ricevere un certificato di malattia proprio perché il medico certifica di aver effettuato una visita ambulatoriale o domiciliare.

Sempre più diffusi sono i servizi di segreteria automatici in cui i pazienti stessi accedono per prenotare la propria visita, richiedere farmaci/impegnative o sottoporre dei quesiti al medico. Se da una parte questo toglie posti di lavoro, dall'altra comporta un costo minore per il titolare.

7) Terapie digitali

Le Terapie Digitali (DTx) rappresentano un'emergente categoria di terapia medica e sono definite come interventi terapeutici basati sulle evidenze scientifiche e rivolte ai pazienti attraverso software che programmano la prevenzione, la gestione e la cura di specifiche condizioni mediche ⁽¹³⁾.

Le patologie per cui sono state attualmente sviluppate comprendono: diabete mellito, disordini psicologici, malattie oncologiche e dipendenze da sostanze.

Le Terapie Digitali sono per natura adatte alla personalizzazione del trattamento.

Quali sono le differenze con la Medicina Digitale di cui parlavamo nel precedente paragrafo? Innanzitutto le DTx utilizzano necessariamente dei software che non vengono sottoposti a iter approvativi al pari di un farmaco o di un dispositivo CE per natura, ma devono dimostrare di produrre risultati in un protocollo di ricerca ben definito e riproducibile. In ultimo devono somministrare una terapia al paziente (da cui il nome).

Torniamo sul concetto di Software as a Medical Device (SaMD).

Abbiamo 3 categorie di Software utilizzate in Sanità: Software come dispositivi medici (SaMD), Software integrato in un dispositivo medico e Software usato nella produzione o nel mantenimento di un dispositivo medico ⁽¹⁴⁾.

L'International Medical Device Regulators Forum (IMDRF) definisce il SaMD come un software utilizzato per uno o più motivi medici che lavora senza essere parte di un hardware di dispositivo medico. ^(15, 16)

I SaMD includono un'ampia gamma di prodotti: app su smartphone, sistemi di supporto decisionale, piattaforme di monitoraggio da remoto e sistemi di registrazione.

Possono essere utilizzati per aiutare il ragionamento clinico, velocizzare la diagnosi, analizzare dati, fare screening, monitorare e allertare medici/pazienti su situazioni di emergenza e perfino gestire patologie croniche.

Il fatto di essere un Software e non un Hardware permette di elicitare l'empowerment dell'utente, che può accedere facilmente e in ogni momento. In questo modo si definisce una maggiore autonomia di utilizzo dell'utente di uno strumento di salute.

Per natura, i SaMD non necessitano di un iter di approvazione dell'FDA, pertanto sono state istituite delle linee guida di utilizzo da parte di FDA e MDR (European Union's medical device regulation) (17, 18, 19). Le SaMD devono sottostare anche al GDPR per quanto concerne la gestione dei dati.

Con le nuove prospettive di IA espresse in precedenza (ML, DL, big data etc.) è facile immaginare quali sono le potenzialità di sviluppo anche dei SaDM.

Le DTx fanno parte della Medicina Digitale e ne rappresenta uno degli ambiti con rischi clinici maggiori, per questo motivo è necessario che lavorino secondo le evidenze scientifiche.

Diverse società si sono espresse sulla soglia di accuratezza necessaria per approvare una DTx: Digital Medicine Society, DTA, HealthXL, NODE e NICE. Nonostante questo si percepisce una certa mancanza di solidità alla base dell'EBM sulle DTx (20).

Ciò è dovuto alla mancanza di un gold standard di riferimento che permetta di valutare gli end-point digitali.

Oltre a questo, le tecnologie in uso durante il trial possono essere modificate e aggiornate.

Anche stabilire un protocollo, degli end-point e dei metodi è ostico in quanto ancora poco esplorato.

La differenza sostanziale con la sperimentazione su farmaco è che la Terapia Digitale mira a produrre un effetto in un organismo vivente senza

far circolare in esso alcuna sostanza. Gli eccipienti e la posologia sono calibrati intorno agli orari e la modalità di somministrazione delle comunicazioni.

La Digital Therapeutics Alliance è un'associazione no-profit fondata allo scopo di sviluppare e promuovere le DTx (22). Ad essa dobbiamo la categorizzazione delle DTx in 4 gruppi:

- 1) Diagnosi di una patologia;
- 2) Prevenzione/gestione di una condizione clinica/patologia;
- 3) Ottimizzazione della terapia;
- 4) Trattamento di una condizione clinica/patologia.

Il campo farmaceutico è fortemente interessato alle DTx, in particolare esse mirano ad aumentare la compliance dei pazienti comunicando e ricordando lo schema di assunzione della terapia tramite notifiche direttamente sul telefono/smartwatch (23).

Un esempio è la Proteus Digital Health® Inc. che ha sviluppato Abilify Mycete™ (24).

Tra le prime applicazioni approvate dall'FDA delle DTx troviamo l'utilizzo nelle dipendenze da sostanze nel 2017 (21).

La prima approvazione di una Terapia Digitale è quindi avvenuta in America, dove attualmente vengono vagliate dal Center for Devices and Radiological Health (CDRH) di FDA dove vengono considerate le app di telefonia. Non tutte le app devono essere approvate prima dell'immissione nel commercio, questo è il caso dei software che non propongono un intervento terapeutico.

In caso contrario devono essere valutate e suddivise per classe di rischio in classe I/II/III. La classe III rappresenta i dispositivi con un profilo di rischio più alto.

La valutazione della classe deve necessariamente precedere l'immissione in commercio.

8) Conclusioni

Le DTx possono essere usate da sole (psicoterapia, consigli comportamentali etc.) o in associazione ad altri trattamenti, può essere considerata una nuova categoria terapeutica ed è in grado di far emergere bisogni non visibili di salute.

Le DTx possono raggiungere persone altrimenti escluse/poco incluse nel sistema di cura.

Rimarrebbero comunque esclusi coloro che non hanno le risorse per avere e utilizzare un dispositivo elettronico (sia per problemi economici, sia per problemi geografici di connessione a Internet, sia per barriere di comprensione come la demenza).

I dispositivi elettronici permettono di raggiungere istantaneamente pazienti allettati o in ambienti rurali e possono interagire con persone portatrici di problemi del sensorio (sordi, ciechi, sordo-muti).

La Medicina Digitale è in grado di utilizzare l'enorme mole di dati prodotta quotidianamente per migliorare la prevenzione, la diagnosi e la cura della popolazione.

Lo sviluppo dell'Intelligenza Artificiale è di valido supporto agli operatori sanitari che utilizzano ingentemente il lavoro intellettuale e il ragionamento.

Le prospettive future di sviluppo promettono grandi traguardi e muovono enormi paure.

Sarebbe auspicabile che gli Stati investano tempo, energie e finanziamenti nella regolamentazione della Medicina Digitale e nell'utilizzo della Tecnologia per la cura della popolazione e per il progresso scientifico.

9) BIBLIOGRAFIA

- 1) Eric Strong, MD¹; Alicia DiGiammarino, MS²; Yingjie Weng, MHS³; et al Chatbot vs Medical Student Performance on Free-Response Clinical Reasoning Examinations JAMA Intern Med. 2023;183(9):1028-1030. doi:10.1001/jamainternmed.2023.2909;
- 2) ARTIFICIAL INTELLIGENCE: DEFINITION, TRENDS, TECHNIQUES, AND CASES Joost N. Kok, Egbert J. W. Boers, Walter A. Kusters, and Peter van der Putten Leiden Institute of Advanced Computer Science, Leiden University, the Netherlands;
- 3) Morley, Floridi, Kinsey et al., 2020;
- 4) Alaiari, Vellino. 2016;
- 5) Floridi, Sanders. 2004;
- 6) Weizenbaum, 1976;
- 7) De Angeli, 2009;
- 8) Spatt, 2016;
- 9) Danaher, 2017;
- 10) Jagatic, Johnson, Jakobsson et al. 2007;
- 11) Artificial Intelligence and life in 2030 – One Hundred Year Study on Artificial Intelligence, Stanford University, 2016;
- 12) <https://www.prnewswire.com/news-releases/global-artificial-intelligence-in-healthcare-market-2019-2025-ai-in-healthcare-to-improve-patient-outcomes---growing-potential-of-ai-based-tools-for-geriatric-patients-300817687.htm>;
- 13) Ji Sun Hong, Chris Wasden, Doug Hyun Han. Introduction of digital therapeutics. Comput methods programs biomed 2021. DOI: 10.1016/j.cmpb.2021.106319;
- 14) <https://www.fda.gov/medical-devices/digital-health-center-excellence/software-medical-device-samd>;

- 15) Software as a Medical Device (SaMD): Clinical Evaluation Guidance for Industry and Food and Drug Administration Staff Document issued on December 8, 2017. The draft of this document was issued on October 14, 2016;
- 16) November 2018: Federal Register Notice [External Link](#)
Disclaimer “Prescription Drug-Use- Related Software” announced the establishment of a docket to seek early input from the public on the agency’s proposed framework for prescription drug-use-related software;
- 17) Software as a Medical Device (SaMD): Key Definitions. IMDRF SaMD Working Group. 9 December 2013;
- 18) "Software as a Medical Device": Possible Framework for Risk Categorization and Corresponding Considerations. IMDRF Software as a Medical Device (SaMD) Working Group. 18 September 2014;
- 19) Software as a Medical Device (SaMD): Application of Quality Management System. IMDRF SaMD Working Group. 2 October 2015;
- 20) Changwon Wang, Chungkeun Lee & Hangsik Shin. Digital therapeutics from bench to bedside. *npj Digit. Med.* 6, 38 (2023).
<https://doi.org/10.1038/s41746-023-00777-z>;
- 21) FDA permits marketing of mobile medical application for substance use disorder. [Accessed 12 February 2019].
<https://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm576087.htm>;
- 22) Digital Therapeutics Alliance. [Accessed 10 January 2019].
<https://www.dtxalliance.org/>;
- 23) Jae-Yong Chung. Digital therapeutics and clinical pharmacology. *Transl Clin Pharmacol.* 2019 Mar;27(1):6-11. English. Published online Mar 27, 2019.
<https://doi.org/10.12793/tcp.2019.27.1.6>
- 24) Voelker R. Digital Pill Gains Approval. *JAMA* 2018;319:14. [doi: 10.1001/jama.2017.19309];

10) RINGRAZIAMENTI

Quando si arriva a questa sezione si percepisce già il sapore agrodolce di una vicenda che volge al termine e, per onestà intellettuale ma anche per umana necessità di emozionarsi, desidero ringraziare le persone che in qualche modo sono state personaggi o comparse di questa storia.

Innanzitutto ringrazio la Curiosità, quel motorino interiore che porta sempre a cercare nuovi stimoli e nuove conoscenze.

Ringrazio la mia Famiglia, che in fondo mi accetta per come sono e mi sostiene in ogni idea più o meno sensata che decido di percorrere.

Ringrazio il Progresso tecnologico e quella fetta di Umanità che si impegna a rendere questo Mondo un posto mediamente accettabile.

Ringrazio la mia Tutor, esempio di eccellenza professionale e disponibilità.

Ringrazio la Bioetica, protagonista indiscussa di questo percorso, che è tanto giovane quanto affascinante.