ISTITUTO ITALIANO di BIOETICA

CONSULTA DI BIOETICA ONLUS

CNB COMITATO NAZIONALE PER LA BIOETICA

CONFERENZA NAZIONALE DI BIOETICA PER LE SCUOLE

L'ETICA DELL'INTELLIGENZA ARTIFICIALE (IA): OPPORTUNITA' E RISCHI

9 aprile 2024

Istituto ospitante – Liceo N. Braucci

Caivano

(Napoli)

XX edizione – anno 2024

ETHICS Commissione per l'Etica e l'Integrità nella Ricerca

La riproduzione e la divulgazione dei contenuti del presente volume sono consentite fatti salvi la citazione della fonte e il rispetto dell'integrità dei dati utilizzati.

Presidenza del Consiglio dei Ministri
COMITATO NAZIONALE PER LA BIOETICA

Conferenza Nazionale di Bioetica per le Scuole
Caivano (NA) 7/8 aprile 2024

Nell'ambito delle manifestazioni di livello nazionale, il Comitato Nazionale per la Bioetica (CNB) ha organizzato, insieme alla Consulta di Bioetica, all'Istituto Italiano di Bioetica e in collaborazione con la Commissione per l'Etica e l'Integrità nella Ricerca del Consiglio Nazionale delle Ricerche (CNR), la Conferenza Nazionale di Bioetica per le Scuole, giunta alla sua ventesima edizione.

L'iniziativa è nata con l'intento di promuovere la riflessione delle giovani generazioni, segnatamente gli alunni delle scuole medie superiori di tutta Italia, sulle tematiche bioetiche.

Sino al 2019 la Conferenza si è tenuta ogni anno in una città diversa, con la partecipazione in presenza di docenti e allievi delle scuole di varie Regioni. Dal 2020, a motivo della pandemia, al fine di non interrompere questo importante appuntamento culturale, si è proceduto in modalità *on line*, utilizzando la piattaforma digitale messa a disposizione dal Centro Interdipartimentale per l'Etica e l'Integrità nella Ricerca del CNR.

Quest'anno si è ritenuto di organizzare l'evento in modalità *blended*, in presenza presso il Liceo Niccolò Braucci di Caivano (Napoli) e con gli altri istituti collegati da remoto.

Il tema prescelto è stato "*L'Etica dell'Intelligenza Artificiale (IA): opportunità e rischi".*

L'evento è stato inaugurato dall'intervento del Presidente del CNB, Prof. Angelo Luigi Vescovi, e dal saluto della Dott.ssa Agnese Camilli, Coordinatrice della Segreteria del CNB.

Il dibattito è stato moderato dalla Prof.ssa Giovanna Razzano, componente del CNB, con gli interventi della Prof.ssa Luisella Battaglia, per l'Istituto Italiano di Bioetica, e del Prof. Lorenzo d'Avack, componente del CNB.

La Prof.ssa Cinzia Caporale, del CNR, e il Prof. Maurizio Mori, della Consulta di Bioetica, hanno concluso i lavori.

Il supporto all'evento è stato garantito dalla Dott.ssa Annalisa Di Mauro, per l'Istituto Italiano di Bioetica; dalla Dott.ssa Marisa Polesana, per la Consulta di Bioetica; dalla Dott.ssa Giorgia Adamo, per il Centro Interdipartimentale per l'Etica e l'Integrità nella Ricerca e la Commissione per l'Etica e l'Integrità nella Ricerca del Consiglio Nazionale per le Ricerche; dalle Dott.sse Monica Bramucci, Raffaella Maria Falco e Lorena Paris per il Comitato Nazionale per la Bioetica.

Seguendo un metodo di lavoro ormai collaudato, gli studenti hanno esposto elaborati individuali e/o di gruppo, divenendo in tal modo autentici protagonisti e autori dell'evento culturale.

Il pregio dei lavori prodotti dagli allievi dei vari istituti, con la preparazione e sotto la guida dei loro docenti, ha indotto l'idea di organizzare il materiale nella presente raccolta, che verrà pubblicata sui siti degli Istituti scolastici e degli Enti organizzatori, quale segno tangibile delle capacità e dell'impegno di tanti ragazzi, che si auspica possano essere d'esempio e di stimolo ai giovani che costruiranno il futuro di questo Paese.

INDICE

Locandina p. 6

Premessa p. 7
Liceo Niccolò Braucci. Caivano (NA)

L'Intelligenza Artificiale: che cos'è e come funziona p. 9
Liceo Niccolò Braucci

L'utilizzo dell'Intelligenza Artificiale in biologia p. 13
Liceo Niccolò Braucci

Relazione sulle implicazioni etiche e giuridiche dell'automazione in medicina
Liceo Niccolò Braucci p. 16

Intelligenza Artificiale nelle biotecnologie: scoperte, progressi e limiti etici. Bioetica dell'I.A. nel campo delle biotecnologie e impatti sulle nostre vite
Liceo Niccolò Braucci p. 19

Esplorazioni etiche nell'era dell'Intelligenza Artificiale: riflessioni e prospettive
Liceo Niccolò Braucci p. 26

Robotica educativa e sue applicazioni nella scuola superiore.
Esperienze di Intelligenza Artificiale simulata a scuola: riflessioni didattiche e neuroetiche
I.I.S. Antonello. Messina p. 38

"Hey Alexa, prenderai il mio posto?" Intelligenza Artificiale e sostenibilità
Liceo G. Galilei. Napoli p. 43

Intelligenza Artificiale e privacy: un incontro possibile?
Liceo P. Gobetti. Genova p. 51

L'era dell'Intelligenza Artificiale tra scienza e società
Liceo Pertini. Genova p. 55

La potenza di controllo dell'Intelligenza Artificiale sulla vita e sulle idee dei cittadini
ITI Marconi. Pontedera p. 59

Evoluzioni dell'I.A. e risvolti etici
Liceo Scientifico e Classico Vieusseux. Imperia p. 77

CONFERENZA NAZIONALE DI BIOETICA PER LE SCUOLE

L'ETICA DELL'INTELLIGENZA ARTIFICIALE (IA): OPPORTUNITÀ E RISCHI

9 aprile 2024 ore 10:00 - 16:00

Ore 10:00 **AVVIO DEI LAVORI**
Introduzione: Luisella Battaglia - CNB **Saluti Istituzionali:** Claudio Mola - Preside Liceo Braucci
Angelo Vescovi - CNB
Agnese Camilli - CNB

Coordina: Giovanna Razzano

Ore 10:45 **Esplorazioni etiche nell'era dell'intelligenza artificiale: riflessioni e prospettive**
Liceo Niccolò Braucci, V DS - Caivano - (NA) - Docente: Giovanna Cennamo
Relatore: Andrea Romano

Ore 11:05 **"Hey, Alexa, prenderai il mio posto?" - Intelligenza artificiale e sostenibilità ambientale**
Liceo Galileo Galilei, IV B - Napoli - Docente: Bruna Celano
Relatori: Simone Corrado, Flavia Venosa

Ore 11:25 **Robotica educativa e sue applicazioni nella scuola superiore: riflessioni neuroetiche e didattiche**
I.I.S. Antonello, IV A ind. grafico - Messina - Docente: Antonio Consolo, Alessandra D'Arrigo, Elisabetta Falduto, Rita Ristagno
Relatori: Sofia Fazio, Laura Ivaldi

Ore 11:45 **Intervento del CNB: Luisella Battaglia**

Ore 12:00 **L'impatto dell'intelligenza artificiale nelle scienze della vita**
Liceo Niccolò Braucci, III CSC - Caivano - (NA) - Docente: Anna Maria Delle Curti
Relatore: D'Ettore Daniele, Serena Ponzo, Luciano Bagnarola

Ore 12:20 **Opportunità e rischi dell'IA: scopriamoli attraverso un debate**
Liceo Gobetti, 2DSE - 2DSU- 1 D/ E /F SU - Genova - Docente: Daniela Malini
Relatori: Martina Torrigino, Giorgia Sbrillo, Gabriele Castagnoli

Ore 12:40 **Intelligenza Artificiale e privacy: un incontro possibile?**
Liceo Gobetti, 5B SU - Genova - Docenti: Alessandra Fabbri, Maria Grazia Fossati
Relatori: Samuele Cosso, Diletta Foli

Ore 13:00 **Intervento del CNB: Lorenzo D'Avack**

Ore 13:15 **PAUSA PRANZO**

Ore 14:00 **RIPRESA LAVORI**

coordina: Maria Pia Garavaglia

Ore 14:00 **Intelligenza artificiale nelle biotecnologie: scoperte, progressi e limiti etici**
Liceo Niccolò Braucci, III DSC - Caivano - (NA) - Docente: Giovanna Cennamo
Relatori: Diletta Dinardo, Alessandra Falco

Ore 14:20 **L'era dell'Intelligenza Artificiale tra scienza e società**
Liceo Pertini, II I SC - Genova - Docente: Rosaria Catanoso
Relatori: Angelica Baldassini, Lorenzo Bernardinelli, Carola Carruba, Ginevra De Paoli, Luca Cuomo, Eleonora Sobrero

Ore 14:40 **Intervento del CNB: Maurizio Mori**

Ore 14:55 **Evoluzioni dell'IA e risvolti etici**
Liceo Viesseux V A - Imperia - Docente: De Meo Raffaella
Relatore: Giulio Rossi

Ore 15:15 **La potenza di controllo dell'IA sulla vita e sulle idee dei cittadini**
ITI G. Marconi, IV BIT ASA - Pontedera (PI) - Docente: Simone Matteoli
Relatori: Alessandro Meloni, Valerio Campana, Razvan Comaniciu

Ore 15:35 **Conclusioni: Alberto Gambino e Maurizio Mori**

Ore 16:00 **FINE LAVORI**

Clicca QUI PER REGISTRARTI ALLA CONFERENZA O INQUADRA IL QR CODE:

segreteria@istitutobioetica.it segreteria@consultadibioetica.org

In collaborazione con:

Premessa

Nella primavera del 2024 il **Liceo "Niccolò Braucci" di Caivano** si è pregiato di ospitare La Conferenza nazionale di Bioetica per le Scuole quest'anno dedicata ai risvolti etici delle applicazioni dell'Intelligenza Artificiale.



Molti sono stati gli spunti emersi dall'articolato dibattito che si è sviluppato tra Studenti ed Esperti e che ha insistito, in particolare, intorno ai nodi di seguito elencati.

É stato, innanzitutto, chiarito il passaggio dalla fase dell'*Intelligenza Artificiale in senso forte* a quella dell'*Intelligenza Artificiale in senso debole*, per il cambiamento di prospettiva entro cui vanno collocate le riflessioni sui risvolti etici di tale innovazione.

Sotto l'onda di ottimismo alimentata dalla straordinaria capacità di archiviazione e dall'altrettanto straordinaria potenza di calcolo raggiunte dagli elaboratori elettronici, nella prima fase, il problema si formulava nei termini "*può una macchina pensare?*" e dunque "*qual è l'ontologia del pensiero?*".

Presto, gli insuccessi collegati all'incontrollabilità del dominio semantico rispetto a quello sintattico e il problema della "computazionalità" algoritmica determinarono la crisi del sogno di McCarthy e Alan Turing, che già era stato quello di Leibnitz, di Pascal, di Raimondo Lullo e cioè: *ogni aspetto dell'intelligenza può essere descritto in termini tanto rigorosi da rendere possibile programmare una macchina in grado di simularli*.

A ridosso degli anni 80 si assistette a una svolta nell'impostazione della ricerca sull'AI: la soluzione di un problema veniva ricercata all'interno di un set di soluzioni note relative a problemi archiviati, attraverso un processo di minimizzazione rispetto a osservazioni raccolte.

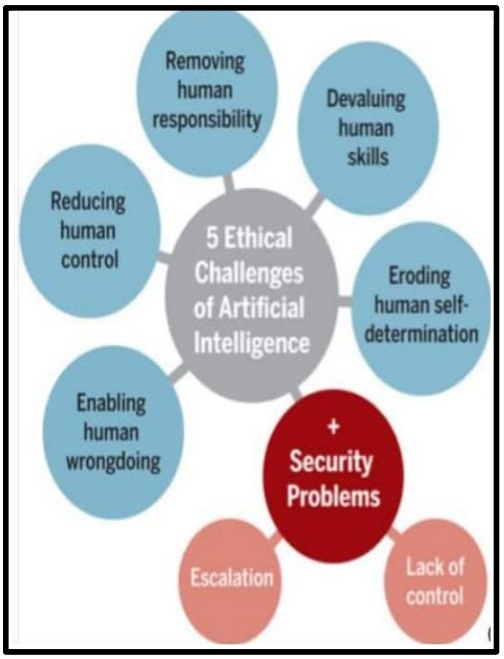
Si tratta di procedure basate sulla conoscenza e quindi sull'elaborazione di una grande mole di informazioni registrate. Era la fase dei sistemi esperti. Si rinunciava al sogno di una logica universale costruita, come tutte le logiche, su un sistema assiomatico e si delimitavano ambiti di intervento per procedure automatizzate attraverso la risorsa di dati disponibili per un processing affidato alle macchine.

Alla logica subentrava la statistica; alle procedure deduttive subentravano quelle induttive implicitamente associate a concetti probabilistici, a verità accettabili con un margine di errore.

Dirompono i problemi collegati:

- Ai dati e, in particolare, alla:
  - rilevazione (nella maggioranza dei casi avviene attraverso la rete e quindi da un universo in qualche modo privilegiato) che può influenzare *direzionalmente* la ricerca della soluzione;
  - conservazione (deve garantire la tutela di quanto attiene alla sfera sensibile dei titolari);
  - elaborazione (deve prevedere la gestione dei *bias cognitivi*, ossia di quelle valutazioni errate della realtà che sono formulate senza "razionalità" e che sono presi per buoni solo perché spontanee o addirittura solo comuni).
- Alla trasparenza delle procedure di elaborazione dei dati.

– Alla responsabilità derivante dalle scelte fatte con il supporto dell'AI (la responsabilità è del fruitore, dello sviluppatore del software, del fornitore dei dati, …).



Un altro problema emerge e coinvolge con meno mediazioni la Scuola e il mondo della formazione.

É evidente che il ricorso all'AI determinerà un abbassamento dell'offerta di lavoro rispetto a mansioni ripetitive che possono essere assolte da computer e programmi, così come è altrettanto evidente che crescerà la richiesta di personale che sia non soltanto esperto nel settore, ma che del settore abbia la capacità di seguire l'evoluzione caratterizzata da forte instabilità, ponendosi a livelli alti di gestione delle procedure.

Le competenze richieste sono quelle che non si consumano in una stagione di fioritura di un prodotto e/o di una tecnologia, per i quali basta un allenamento, ma quelle che sanno seguire le trasformazioni, che hanno potere sintetico e congiuntamente acutezza analitica, che evidenziano capacità di interrogare e di sentire il contesto per prevederne le svolte evolutive e anticipare sulle soluzioni. In altri termini, le competenze richieste non sono settoriali ma congiuntamente specialistiche e molteplici per affrontare il mondo della complessità.

L’INTELLIGENZA ARTIFICIALE: CHE COS’È E COME FUNZIONA

Daniele D’Ettore, Luigi Falibretti - Classe III Csc

1. Definizione e classificazione dell'Intelligenza Artificiale

L'intelligenza artificiale (IA) rappresenta il nucleo tecnologico che permette di emulare i processi dell'intelligenza umana attraverso l'implementazione di algoritmi in un ambiente computazionale dinamico. In altre parole, l'obiettivo principale dell'IA è quello di sviluppare sistemi informatici in grado di pensare ed agire in modo simile agli esseri umani. Attualmente, l'IA può essere suddivisa in due tipologie distinte:

- l'IA debole (o weak AI) che si fonda sul “come se” ovvero agisce e pensa come se avesse un cervello. La meta dell’IA debole non è quella di realizzare macchine che abbiano un’intelligenza umana, quanto, invece, sistemi che possano agire con successo in alcune funzioni complesse umane, come, ad esempio, la traduzione automatica di testi. In un ipotetico confronto tra uomo e macchina, la mente umana mantiene la sua supremazia. La macchina non è capace di pensare in maniera autonoma, svolge egregiamente il suo compito ma ha bisogno della presenza dell’uomo.
- l'IA forte (o strong AI) si basa su un processo di ricerca mirato a replicare il più fedelmente possibile la fisiologia del cervello umano; in questo caso, la macchina non è soltanto uno strumento ma diventa essa stessa una mente con una capacità cognitiva non distinguibile da quella umana.

2. Genesi e importanza dell'Intelligenza Artificiale

Quando si parla di intelligenza artificiale spesso si pensa a tecnologie futuristiche, a robot capaci di ragionare e prendere decisioni autonomamente. Tuttavia, l'IA e le sue applicazioni sono molto più concrete e diffuse di quanto si possa immaginare.

Oggi, l'IA è utilizzata in diversi settori della vita quotidiana seppur con un impatto meno invasivo rispetto a quanto, talvolta, rappresentato nei film di fantascienza. È interessante sottolineare come i progressi scientifici sul tema dell’intelligenza artificiale siano partiti non dal campo dell’informatica ma da quello della biologia. In proposito, nel 1969 studenti e ricercatori del *Carnegie Institute of Technology* svilupparono un programma chiamato DENDRAL, capace di ricostruire molecole semplici basandosi sui dati ottenuti da uno spettrometro di massa. Questo fu un passo rivoluzionario che diede avvio ad un diverso modo di intendere l’intelligenza artificiale, focalizzando, in tal senso, l’attenzione sui sistemi esperti, capaci di fornire soluzioni specifiche tenendo come punto di riferimento il caso concreto.

Sebbene alcune forme di IA esistano da oltre mezzo secolo, i progressi nella potenza di calcolo, l'abbondanza di dati e lo sviluppo di nuovi algoritmi hanno portato a significativi avanzamenti tecnologici solo negli ultimi anni. Le future applicazioni potrebbero portare a cambiamenti di ampia portata, nonostante l'IA sia già parte integrante delle nostre vite quotidiane.

3. Funzionamento dell'IA

Alla base del funzionamento dell'IA si trovano gli algoritmi, ovvero sequenze di istruzioni che definiscono le operazioni eseguite dalla macchina sui dati per ottenere i risultati desiderati. Gli algoritmi integrano anche espressioni matematiche, utili nella risoluzione di problemi di vario genere.

Gli algoritmi vengono 'addestrati' utilizzando esempi umani, fornendo alla macchina la base per imparare. Questi ultimi sono suddivisi in macrocategorie, ognuna delle quali contenente diverse procedure mirate alla risoluzione di un problema specifico. Le principali aree algoritmiche includono:

- Machine Learning (Apprendimento Automatico): questa tecnologia consente alle macchine di apprendere dai dati, migliorando le loro prestazioni nel tempo. Attraverso il Machine Learning i sistemi di intelligenza artificiale possono essere utilizzati per classificare immagini, prevedere dati e rilevare anomalie. Esistono quattro modalità principali di apprendimento:

1. apprendimento supervisionato: in questo caso il compito assegnato alla macchina è quello di trovare il miglior modo di svolgere una funzione sulla base di valori di partenza (dati di addestramento) assegnati dall'uomo al momento della sua creazione. La macchina, sotto la supervisione di un essere umano, tenta la risoluzione di un problema fino al raggiungimento del risultato desiderato; in questo modo, sarà poi in grado di risolvere non soltanto il quesito iniziale ma anche richieste affini;
2. Apprendimento non supervisionato: in questo caso la macchina utilizza dati non etichettati al fine di scoprire informazioni riguardo la funzione che dovrà eseguire. Partendo da questi dati, individua dei modelli che aiutano a risolvere problemi di clustering o di associazione. Questo è particolarmente utile quando gli esperti non sono sicuri delle proprietà comuni all'interno di un set di dati;
3. apprendimento semi-supervisionato: si tratta di un ibrido tra i primi due; si verifica quando solo una parte dei dati di input forniti è stata etichettata. Per poter affermare il corretto apprendimento della macchina si richiede che l'output sia sufficiente a completare la funzione;
4. apprendimento forzato: la macchina viene posta in un ambiente sconosciuto senza alcun tipo di dato e viene lasciata libera di muoversi autonomamente. Ogni minimo output trasmette un feedback che viene studiato per capire il livello di apprendimento della macchina. Alla fine del percorso la macchina svolgerà funzioni varie autonomamente.

- Il Natural Language Processing (NLP) è una tecnologia rivoluzionaria che permette ai sistemi di intelligenza artificiale di comprendere ed interpretare il linguaggio umano naturale. Grazie a questa innovazione, le macchine possono elaborare testi e discorsi, rispondere a domande e persino analizzare i sentimenti espressi nel linguaggio. Le macchine addestrate con il NLP non solo imparano il linguaggio umano, sia verbale che corporeo, ma sono in grado di replicarlo autonomamente. Possono condurre ricerche, spiegare concetti complessi e rispondere alle richieste degli utenti in modo preciso ed efficiente.
- La Computer Vision: consente ai sistemi di intelligenza artificiale di elaborare e analizzare immagini e video in maniera straordinariamente accurata. Questa tecnologia abilita il riconoscimento di oggetti, persone e situazioni, oltre all'analisi dettagliata di immagini mediche. La Computer Vision, inoltre, va al di là del semplice riconoscimento e consente di creare nuove immagini, video e altro partendo da zero, basandosi su input testuali forniti dagli utenti. È un'abilità incredibile che negli ultimi periodi, con IA come DALL-E2 o il modello Sora (Text-to-video) di OpenAI, ha ampliato notevolmente le capacità creative e diagnostiche delle macchine intelligenti.

Esiste una diffusa preoccupazione riguardo alla possibilità che l'autonomia delle macchine possa sfuggire al controllo umano. Tuttavia, è importante sottolineare che l'uomo ha ancora un ruolo fondamentale nel processo di addestramento degli algoritmi; quindi, parlare di una reale autonomia decisionale delle macchine potrebbe essere prematuro. Nonostante le capacità logiche dell'IA siano alla base di discussioni prolungate, è importante riconoscere che, attualmente, l'IA esegue principalmente correlazioni tra dati precedentemente immagazzinati piuttosto che processi di ragionamento deduttivo simili a

quelli umani. Un'area di particolare interesse è il deep learning, che cerca di simulare il funzionamento del cervello umano, anche se, spesso, restituisce risultati di difficile interpretazione, noti come "black box". Oltre queste incertezze tecniche, c'è anche la preoccupazione che l'IA possa superare l'uomo nel suggerire decisioni, portando a una possibile riduzione delle competenze umane, fenomeno noto come "deskilling".

In conclusione, alla luce di quanto esposto, è possibile affermare che, seppure l’intelligenza artificiale sia in continua evoluzione, resta essenziale mantenere il contributo umano nell'addestramento e nell'utilizzo di tali sistemi.

Fonti

1. *Articolo del 12/03/2024 - Redazione osservatori digital innovation* - Politecnico di Milano – Department of Management Engineering // www.osservatori.net/it/assistenza/cookie-center
2. *"Che cos'é 'l'intelligenza artificiale"* - NetApp // https://www.netapp.com/it/artificial-intelligence/what-is-artificial-intelligence/
3. *"Che cos'è l'intelligenza artificiale"*-SAS,Analytics Software Soluzioni // *www.sas.com/it_it/insights/analytics/what-is-artificial-intelligence.html*
4. *Articolo del 04/12/23 - "Metodi di addestramento algoritmi IA"* - Avv. Luisa Di Giacomo per diritto.it // *www.diritto.it/metodi-addestramento-algoritmi-ia-garante-indaga/*
5. *"Intelligenza Artificiale: Cos'è, come funziona e a cosa serve"* - I.A., Il portale dedicato all'intelligenza artificiale // www.intelligenzaartificiale.it/

L'UTILIZZO DELL'INTELLIGENZA ARTIFICIALE IN BIOLOGIA

Serena Ponzosollo, Alfonsina Fatima, Sabrina Bernardo - Classe III Csc

1. Affidabilità dell'IA in biologia

L'intelligenza artificiale viene utilizzata in diversi settori della biologia, tra cui la bioinformatica, la modellazione molecolare, la diagnosi medica e la scoperta dei farmaci. Tuttavia, anche l'IA ha un margine di errore e, pertanto, non può essere considerata sempre attendibile. Il margine di affidabilità dipende da fattori di vario genere.

L'IA riesce a essere sempre più precisa ed efficiente avendo a disposizione una larga scala di dati, ma se questi ultimi dovessero risultare compromessi, imprecisi, o non rappresentativi,i risultati potrebbero essere inaccurati o distorti. Ad esempio, nell'analisi genomica, la completezza e l'accuratezza delle sequenze genomiche influenzano direttamente le previsionidegli algoritmi dell'IA. Inoltre, è essenziale valutare la dimensione del campione e l'eterogeneità dei dati: modelli addestrati su dataset troppo piccoli o non rappresentativi possono essere soggetti a over fitting e, di conseguenza, produrre previsioni poco affidabili.

Gli algoritmi utilizzati devono essere accurati, generalizzabili e robusti per poter far fronte a possibili variazioni nei dati di input. Devono, inoltre, essere in grado di gestire dati biologicicomplessi e variabili. Un esempio può essere l'analisi di immagini mediche in cui c'è bisognodi saper riconoscere e interpretare variazioni sottili nei tessuti o nelle strutture cellulari. A questo scopo vengono utilizzati algoritmi di deep learning per il riconoscimento di pattern complessi. Infine, gli algoritmi, al fine di garantire la loro accuratezza e robustezza su una larga scala di dati devono essere validati.

La validazione dei modelli di IA è un passaggio critico e ha lo scopo di valutarne l'affidabilità.Questo può essere effettuato utilizzando dati di test indipendenti o attraverso tecniche come,ad esempio, la cross-validazione. È importante che i modelli siano in grado di generalizzare ben oltre i dati di addestramento per essere considerati affidabili.

La verifica dei risultati è altrettanto importante. Ad esempio, nel caso delle predizioni di strutture proteiche, i risultati dell'IA possono essere confrontati con esperimenti di cristallografia a raggi X o risonanza magnetica nucleare (RMN) per confermare la precisione delle previsioni.

Modelli troppo complessi o black box possono risultare difficili da interpretare e tanto comporterebbe una limitazione di utilizzo limitandone da parte degli esperti del settore. È necessario, quindi, che le previsioni dell'IA siano interpretabili e comprensibili, tanto in quanto non è la soluzione universale a tutti i problemi biologici e non può sostituire completamente gli esperti essendo soltanto un mezzo complementare alle metodologie tradizionali. Agli esperti deve essere resa semplice la comprensione e il confronto tra i risultati forniti dall'IA e quelli ottenuti attraverso metodi di ricerca e analisi convenzionali.

2. Consenso e autonomia

Quando si tratta di dati personali sensibili, forniti all'IA per svolgere studi medici o biologici,bisogna tenere in considerazione due aspetti fondamentali.

In primo luogo, nell'ambito della ricerca biologica, gioca un ruolo fondamentale il consenso al trattamento dei dati personali. È necessario che le persone coinvolte nello studio o nel campionato dei dati biologici prestino il consenso informato al trattamento dei dati rientranti nella sfera dei cd. dati sensibili. Per consenso si intende il permesso o l'autorizzazione fornitadalle persone interessate – nonché titolari di tali dati – prima che vengano

utilizzati o elaborati. Questo è particolarmente importante in quanto le informazioni genetiche o dati medici rientrano nella categoria dei dati sensibili.

Anche quando si utilizzano dati già esistenti, come quelli provenienti da banche dati genomiche o da studi clinici, è importante rispettare le politiche di consenso stabilite dai fornitori dei dati e garantire che vengano seguite le normative sulla privacy e sulla protezione dei dati personali.

L'IA deve necessariamente rispettare i principi etici e legali nel trattamento dei dati biologici, come la privacy, la non discriminazione e il rispetto della dignità umana. Gli algoritmi dovrebbero essere progettati per proteggere la privacy dei dati e per evitare l'uso improprio odiscriminatorio delle informazioni biologiche.

In secondo luogo, bisogna concentrare l'attenzione sull'autonomia che, invece, si riferisce alla capacità dell'IA di elaborare i dati in modo indipendente, senza la supervisione umana diretta. Questo è particolarmente rilevante quando si utilizzano tecniche di apprendimento automatico o di apprendimento profondo che possono analizzare grandi quantità di dati in modo autonomo al fine di identificare pattern o fare previsioni.

L'autonomia dell'IA non significa che essa agisce senza controllo o senza la supervisione umana; gli algoritmi di IA devono necessariamente essere sviluppati ed addestrati da esperti umani; i risultati, invece, devono essere valutati e interpretati da ricercatori esperti del settore.

Riferimenti bibliografici

1. M.C. Schatz, *Biological data sciences in genome research*., 2012.
2. *Opportunities and obstacles for deep learning in biology and medicine*, 4 April 2018.
3. Alphafold, *Come l'IA rivoluziona la biologia*, 17 Agosto 2022.
4. *Algoritmi e diritti umani: qual è il punto d'incontro?* 12 Gennaio 2021.
5. European Parliament and Council (2016) Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016.
6. *The ethics of algorithms: Mapping the debate*. Big Data & Society, 2016.
7. *Privacy in the age of medical big data*. Nature Medicine, 2019.

RELAZIONE SULLE IMPLICAZIONI ETICHE E GIURIDICHE DELL'AUTOMAZIONE IN MEDICINA

Luciano Bagnarola, Simone Prota - Classe III CSC

L'automazione, in campo medico, rappresenta una rivoluzione copernicana che promette di migliorare significativamente l'accuratezza delle pratiche mediche e di ridurre gli incidenti, contribuendo, in tal senso, alla diminuzione della mortalità e, più in generale, ad un miglioramento delle prestazioni sanitarie.

Sempre più frequentemente, infatti, si parla, in ambito ospedaliero, del ricorso a macchinari che possano sostituirsi all'operatore professionale, sia in fase esecutiva che diagnostica. In proposito, dunque, i sistemi di intelligenza artificiale sono destinati a intervenire sul processo di cura sia in fase diagnostica che terapeutica, permettendo all'operatore sanitario di pervenire, da un lato, a diagnosi più certe e precise - es. tramite *imaging* radiologico, capace di analizzare lastre e referti con una percentuale di esattezza e precisione inimmaginabili a occhio umano - dall'altro a terapie chirurgiche più efficaci e meno invasive grazie all'utilizzo della robotica[1].

Si evidenzia, infatti, che gli indubbi benefici apportati, in termini di tutela del diritto alla salute del paziente, quali diagnosi precoci e somministrazione di terapie sempre più precise e meno invasive, vengono conclamati dal Parlamento europeo che, con la Risoluzione del 16 febbraio 2017, sottolinea l'importanza di incentivare lo sviluppo di tali tecnologie[2].

Se, da un lato, l'utilizzo dell'intelligenza artificiale in medicina permette di pervenire a notevoli vantaggi per la salute del paziente, dall'altro è innegabile che la mancanza di una disciplina di regolazione *ad hoc* della materia comporti per il sanitario il rischio di incorrere facilmente in responsabilità civile[3].

Ne deriva una rivoluzione senza precedenti nell'ambito della relazione di cura medico-paziente in cui il medico, che nasce come *pater* ed indiscusso mentore del paziente nel modello paternalistico, si evolve ad alleato nella gestione della terapia del paziente nel modello della reciprocità; rapporto destinato ancora una volta a mutare nel momento in cui nella relazione suddetta interviene un terzo elemento: l'intelligenza artificiale. La conseguenza diretta del sempre maggior uso di tecnologica automatica in ambito sanitario è una nuova concezione del medico come *human in command*, che mantiene in alcuni casi il solo ruolo di controllo del risultato finale della macchina (output) e di validazione del medesimo[4].

A fronte, quindi, di vantaggi in termini di assistenza più personalizzata e tempestiva ai pazienti, l'implementazione di sistemi di intelligenza artificiale (IA) in ambito medico non è priva di rischi e solleva complessi problemi sia etici che giuridici.

In particolare, la questione della responsabilità emerge come uno dei nodi cruciali nell'utilizzo e nello sviluppo dei sistemi di IA in medicina. È fondamentale stabilire chi debba essere considerato responsabile per le decisioni prese attraverso questi sistemi, che possono essere influenzate da fattori come la programmazione errata delle macchine o il

---

[1] *Analisi delle implicazioni etiche e giuridiche dell'automazione in medicina.* CNB (Comitato Nazionale per la Bioetica).

[2] *"Raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica (2015/2103(INL))*", in *Gazzetta Ufficiale dell'Unione europea*.

[3] L. Buonanno, La responsabilità civile nell'era delle nuove tecnologie: l'influenza della blockchain, in Resp. civ. prev., 2020.

[4] U. Ruffolo, *L'Intelligenza artificiale in sanità: dispositivi medici, responsabilità e "potenziamento"*, in Giur. it., 2021, 502.

loro cattivo impiego da parte degli operatori sanitari. Questo solleva interrogativi sui ruoli e sulle responsabilità dei vari attori coinvolti, tra cui i progettisti, i venditori del *software,* gli utilizzatori (come i medici) e altre parti interessate. È, quindi, necessario un chiarimento a livello legislativo e politico per definire l'attribuzione della responsabilità e tracciare una catena di responsabilità analoga a quella degli errori medici tradizionali. In altri termini, l'introduzione di tecnologie automatizzate in medicina richiede un intervento normativo che sia in grado di regolare le innovazioni nel settore in modo da garantire la tutela del paziente e stabilire la responsabilità dei vari attori coinvolti.La complessità della catena di responsabilità risiede nella distinzione tra l'IA come creatore dell'algoritmo e il produttore, il venditore e l'utilizzatore del software. Questa situazione sottolinea l'importanza di una riflessione interdisciplinare che coinvolga sia le scienze giuridiche che quelle mediche per delineare le responsabilità connesse all'utilizzo dell'IA in medicina e garantire una distribuzione equa e trasparente delle stesse.
Le raccomandazioni del Comitato Nazionale di Bioetica (CNB) evidenziano l'importanza di adottare precauzioni nell'utilizzo dell'IA in medicina. Si suggerisce di predisporre controlli accurati per l'addestramento delle macchine e di condurre sperimentazioni adeguate per garantire sicurezza ed efficacia nell'uso di queste nuove tecnologie. Inoltre, si raccomanda di informare correttamente i pazienti sui rischi e benefici dell'utilizzo dell'IA, garantendo un adeguato controllo umano nelle interazioni tra uomo e macchina. È fondamentale anche ripensare alla formazione dei professionisti della salute per adattarli al cambiamento tecnologico e introdurre principi etici nei codici di condotta e nella formazione degli ingegneri e sviluppatori. Risulta necessario promuovere una consapevolezza pubblica delle opportunità e dei rischi delle nuove tecnologie al fine di sollecitare un aggiornamento normativo sulla responsabilità nell'applicazione dell'IA in medicina, con particolare riguardo alla ricerca pubblica nel settore sanitario.

Si tratta di un contesto che impone una nuova e profonda riflessione giuridica. Del resto, ad ogni generazione di giuristi tocca il compito di misurare l'impatto della tecnologia sulle categorie ordinanti del diritto privato. Il giurista del primo '900 si è trovato a dover registrare i mutamenti che la motorizzazione del traffico stradale ha provocato sulla responsabilità civile, quello del secondo '900 ha dovuto fare i conti con le conseguenze che la diffusione di modelli riproduttivi di massa ha generato sul contratto; alla generazione di giuristi dell'ultima porzione del secolo scorso è toccato indagare il ruolo della digitalizzazione telematica sul contratto e sul diritto d'autore. Al giurista odierno spetta, forse, l'onere maggiore, ossia quello di neutralizzare l'impatto che l'intelligenza artificiale sembra in grado di determinare su ogni ramo del diritto, con un livello di pervasività senza precedenti. La struttura della responsabilità civile spesso appare, facendo un confronto con gli strumenti dell'intelligenza artificiale, inadatta a rispondere efficacemente ai problemi che l'uso della robotica pone, specie se volgiamo lo sguardo all'ambito sanitario. Siamo di fronte alla nascita di nuovi sistemi di responsabilità, la cd. responsabilità civile algoritmica, o responsabilità robotica[5].

In conclusione, è opportuna la necessità di una stretta cooperazione tra intelligenza umana e intelligenza artificiale nella somministrazione della prestazione di cura, con una consapevolezza di fondo: "Nessun robot, nessun algoritmo potrà mai sostituire il medico, perché l'arte professionale sta nell'interpretare i dati secondo le esigenze del paziente"[6].

[5] GIURETA Rivista di Diritto dell'Economia, dei Trasporti e dell'Ambiente, XIX 2021.
[6] F. Anelli, Presidente della FNOMCeO, in occasione degli *Stati generali della professione medica* del 16 e 17 maggio 2019, Roma.

INTELLIGENZA ARTIFICIALE NELLE BIOTECNOLOGIE:SCOPERTE, PROGRESSI E LIMITI ETICI.
BIOETICA DELL'INTELLIGENZA ARTIFICIALE NEL CAMPO DELLE BIOTECNOLOGIE E IMPATTI SULLE NOSTRE VITE

Alessandra Falco, Diletta Dinardo - III Dsc

1. Presentazione

Quando si parla di Intelligenza Artificiale, si pensa subito a tecnologie all'avanguardia, a robot in grado di comprendere e decidere le azioni da compiere e di un mondo futuristico in cui macchine e uomini convivono. In realtà, l'Intelligenza Artificiale e il suo utilizzo sono molto più reali di quanto si possa immaginare e vengono oggi utilizzati in diversi settori della vita quotidiana. Si tratta tuttavia di utilizzi meno invasivi di quello che si pensa o di quanto viene mostrato spesso dai film di fantascienza che hanno trovato, nel tema dell'Intelligenza Artificiale, lo spunto per molte serie più o meno di successo.L'intelligenza artificiale supporta già gran parte delle nostre attività quotidiane e ci aiuta a prendere decisioni più informate. Il suo punto di forza è la capacità di derivare conoscenza da grandi quantità di dati.

2. Definizione di IA e recenti sviluppi

In termini tecnici, l'Intelligenza Artificiale è un ramo dell'informatica che permette la programmazione e progettazione di sistemi hardware e software che permettono di dotare le macchine di determinate caratteristiche che vengono considerate tipicamente umane quali, ad esempio, le percezioni visive, spazio-temporali e decisionali. Si tratta cioè, non solo di intelligenza intesa come capacità di calcolo o di conoscenza di dati astratti, ma anche e soprattutto di tutte quelle differenti forme di intelligenza che sono riconosciute dalla teoria di Gardner, e che vanno dall'intelligenza spaziale a quella sociale, da quella cinestetica a quella introspettiva.

Per come viene definita oggi, l'Intelligenza Artificiale nasce con l'avvento dei computer e la sua data di nascita viene fissata come il 1956. Proprio in quest'anno, infatti, si parlò per la prima volta di Intelligenza Artificiale durante un convegno che si tenne in America e che vide la partecipazione di alcuni dei più importanti ricercatori dell'ambiente. Durante questo storico convegno, furono presentati alcuni programmi già capaci di effettuare alcuni ragionamenti logici, in particolar modo legati alla matematica. Il programma Logic Theorist, sviluppato da due ricercatori informatici, Allen Newell e Herbert Simon, era infatti in grado di dimostrare alcuni teoremi di matematica partendo da determinate informazioni.

Gli anni successivi alla nascita dell'Intelligenza Artificiale furono anni di grande fermento intellettuale e sperimentale: università e aziende informatiche, in particolare l'IBM, puntarono alla ricerca e allo sviluppo di nuovi programmi e software in grado di pensare e agire in determinati aspetti come gli esseri umani. Nacquero così programmi in grado di dimostrare teoremi sempre più complessi e, soprattutto, nacque il Lisp, ossia il primo linguaggio di programmazione che per oltre trent'anni fu alla base dei software di Intelligenza Artificiale. La particolarità degli anni Cinquanta-Sessanta fu soprattutto il sentimento di ottimismo che sosteneva tutte le ricerche e le sperimentazioni relative a questo ramo: tuttavia, se da un lato si riuscirono a sviluppare software sempre più sofisticati e in grado di risolvere soprattutto elaborazioni matematiche, dall'altro si iniziarono a vedere le prime limitazioni dell'Intelligenza Artificiale, che non sembrava poter riprodurre le capacità intuitive e di ragionamento tipiche degli esseri umani. L'avventura scientifica dell'AI, nata più di 60 anni fa, include due principali aree di ricerca che in realtà sono spesso combinate.

Una è spesso chiamata l'AI simbolica o logica. In questa modalità, i ricercatori e gli

sviluppatori capiscono come definire un problema (tramite simboli che descrivono le sue componenti principali) e come risolverlo, poi codificano questa soluzione in un algoritmo o delle regole, e trasformano questa soluzione in codice software per comunicarlo alla macchina. Questo modo di costruire macchine "intelligenti" è adatto quando il problema è ben definito.

Un'altra grande area di ricerca è invece basata sull'analisi di dati e l'apprendimento automatico (machine learning), in cui l'AI riesce a trovare correlazioni e pattern invisibili alle persone perché "nascosti" in grandi quantità di informazioni. Questo approccio funziona molto meglio di quello basato su regole e logica quando il problema da risolvere è troppo vago o ha troppe possibili istanze per poter essere descritto precisamente.

Mentre le tecniche di AI basate sul ragionamento logico hanno bisogno di un ambiente "controllato" in cui le macchine si trovano a funzionare, le tecniche di apprendimento automatico hanno dato alle macchine capacità "percettive", quali la possibilità di interpretare correttamente un testo, un suono, una immagine, o un video. Questo ha permesso di ampliare enormemente il campo applicativo dell'AI negli ultimi anni.

La storia dell'IA si è così sviluppata con una serie di alti e bassi nelle cosiddette "AI seasons", dove a giorni di primavera (con molte scoperte e progetti attivi) seguirono giorni d'inverno (dove le scoperte erano quasi nulle e la delusione per i risultati molto profonda).

Oggi, stiamo vivendo un'estate molto soleggiata, con un rinnovato interesse per i metodi di IA e applicazioni mai viste prima. Ciò è essenzialmente dovuto alla disponibilità di tre risorse principali: nuove metodologie per la costruzione e il perfezionamento dei modelli di IA (da ricerche basilari in informatica e nelle relative discipline), un'enorme disponibilità di dati di vario tipo (testi, immagini, suoni, dati strutturati, ecc.) e infine risorse computazionali su larga scala e ad alta prestazione.

Lo sviluppo del cosiddetto approccio deep learning ci permette di affrontare problemi molto difficili che richiedono capacità complesse di ragionamento.

L'impatto di tali metodologie spinge le attuali tecnologie verso nuove applicazioni, producendo effetti collaterali complessi sull'ambiente quotidiano, su quello socioeconomico, giuridico ed etico.

3. Sviluppi dell' IA nel campo delle biotecnologie

Il potenziale impatto dell'intelligenza artificiale nelle biotecnologie è enorme. Sfruttando gli algoritmi di apprendimento automatico e altre tecnologie di intelligenza artificiale, i ricercatori possono analizzare set di dati complessi in modo estremamente efficiente ed efficace. Ciò ha il potenziale per sbloccare nuove scoperte, accelerare il processo di sviluppo dei farmaci e migliorare i risultati dei pazienti affetti da un'ampia gamma di malattie e condizioni. L'intelligenza artificiale sta trasformando la biotecnologia, dalla scoperta di farmaci alla genomica all'assistenza sanitaria. Gli algoritmi di apprendimento automatico e altre tecnologie di intelligenza artificiale consentono un'analisi più rapida e accurata di set di dati complessi. Un esempio delle enormi possibilità di sviluppo della Intelligenza Artificiale è la modellistica molecolare, ovvero una tecnica chiave nella ricerca biotecnologica che consente di comprendere e persino prevedere la struttura e la funzione di molecole biologiche, come le proteine. Il software AlphaFold, sviluppato da DeepMind e dall'EMBL-EBI, è un ottimo esempio dell'uso dell'IA nell'ambito delle biotecnologie. Utilizzando le reti neurali, AlphaFold è in grado di prevedere la struttura delle proteine con una grande precisione. In particolare, questo programma si è distinto nella previsione di strutture per le quali non esistevano modelli precedenti. Ciò è utile per la progettazione di farmaci, poiché la comprensione della struttura tridimensionale delle proteine bersaglio consente di identificare composti che si legano in modo specifico ed efficace. Ciononostante, gli sviluppatori sono consapevoli di dover continuare a migliorare questo strumento, soprattutto nell'area della previsione del ripiegamento delle proteine. Attualmente, sono in corso

numerosi studi per testare nuove molecole efficaci contro malattie come il cancro, la resistenza agli antibiotici e l'Alzheimer. Tuttavia, i tradizionali metodi di test spesso sono lenti, costosi e hanno un basso tasso di successo. Fortunatamente, la biotecnologia ha trovato un prezioso alleato nell'intelligenza artificiale. L'IA non sostituisce i biotecnologi, ma agisce come un assistente intelligente che accelera il processo di ricerca e di screening di composti promettenti.

La genetica rappresenta un aspetto essenziale delle biotecnologie ed è una delle discipline che trae maggior vantaggio dall'IA. Grazie agli algoritmi di machine learning, è possibile analizzare ampie collezioni di dati genomici e proteomici per identificare geni ed enzimi rilevanti per la sintesi di prodotti specifici. Questo processo consente di ottimizzare le vie metaboliche e di manipolare i microrganismi al fine di ottenere una produzione più efficiente di biocarburanti, prodotti chimici industriali e farmaci. Zymergen utilizza l'intelligenza artificiale per accelerare l'ingegneria genetica e la modifica dei microrganismi. Utilizzando algoritmi di machine learning, analizzano e progettano sequenze genetiche per migliorare la produzione di metaboliti di interesse. L'intelligenza artificiale diventa così un potente strumento per progettare microrganismi geneticamente modificati in grado di produrre metaboliti in modo più efficiente.

L'applicazione dell'intelligenza artificiale nelle biotecnologie agricole può fornire soluzioni ai problemi causati dal riscaldamento globale. Ciò include l'identificazione di colture resistenti ai cambiamenti ambientali e agli eventi estremi, come i periodi di siccità. L'intelligenza artificiale e l'analisi delle immagini consentono inoltre di individuare precocemente malattie e parassiti delle piante, facilitando così l'applicazione accurata e automatizzata di agenti fitosanitari. In breve, l'integrazione dell'IA in agricoltura offre un grande potenziale per migliorare la produttività, la sostenibilità e la sicurezza alimentare in un mondo in continua evoluzione.

L'industria biotecnologica si basa fortemente sull'elaborazione, filtraggio, analisi e condivisione efficiente dei dati. Aziende biotecnologiche e organizzazioni sanitarie in tutto il mondo gestiscono grandi quantità di informazioni. Queste includono: dati per la produzione di farmaci, analisi chimiche di composti diversi, sequenziamento di RNA e DNA, studi sugli enzimi e altri processi biologici simili. Per affrontare queste sfide in modo rapido ed efficiente, l'utilizzo di soluzioni software basate sull'intelligenza artificiale è fondamentale. Queste soluzioni consentono di accelerare il processo di lavoro e ridurre gli errori manuali, contribuendo così a un progresso più rapido e preciso nel campo della biotecnologia. In breve, l'intelligenza artificiale ha dimostrato di essere uno strumento prezioso nel campo delle biotecnologie, stimolando la ricerca e aprendo nuove possibilità. Inoltre, combinando l'intelligenza umana con le capacità analitiche dell'IA, possiamo compiere progressi significativi in settori quali la scoperta di farmaci, la modellazione molecolare e l'analisi dei dati.

Il termine "Biotecnologie Molecolari" (MB) copre tutte queste tecnologie di editing diretto sul genoma e le loro applicazioni. Le applicazioni spaziano dall'energia, all'ambiente, all'alimentazione e all'agricoltura. Includono anche applicazioni per la salute umana che fanno uso di DNA, genomi e organismi non umani alterati per scopi di salute umana, nonché interventi diretti sul genoma umano come la terapia genica, interventi sulla linea germinale genetica umana e miglioramento genetico umano.

Le applicazioni infantili del MB includono la riprogettazione del genoma del lievito e del suo metabolismo per consentire la produzione di artemisinina, una sostanza usata per trattare la malaria, e la creazione di compartimenti sintetici di DNA batterico per rilevare la malattia nel corpo umano Nel corso della sua argomentazione, l'EGE fa riferimento al dibattito etico-filosofico tra posizioni antropocentriche ed ecocentriche in etica e sostiene che anche un approccio ecocentrico non deve necessariamente portare a un'opposizione radicale alla biologia sintetica, ma può consentire il commercio. La sezione sulla biosicurezza introduce e spiega anche il principio di precauzione. Si sottolinea che questo

principio non favorisce semplicemente lo status quo, poiché la passività può comportare anche dei rischi. Tuttavia, questo principio porta l'EGE a richiedere, tra le altre cose, studi a lungo termine sull'impatto degli organismi sintetici sull'ambiente prima che tali organismi siano autorizzati per l'uso non confinato.

Oltre alla sua affinità per le questioni etico-filosofiche della biologia sintetica e all'attenzione giuridica, che emerge quando l'EGE discute della proprietà intellettuale, l'EGE è vicino alle posizioni di conservazione della natura e di critica tecnologica, ma allo stesso tempo si sforza di dimostrare la compatibilità generale di queste posizioni con il progresso tecnologico. Adottando i diritti umani come fondamento delle linee guida, l'HLEG opta esplicitamente per una prospettiva "antropocentrica". L'AI HLEG basa le sue linee guida etiche sulla dignità umana, vale a dire sulla convinzione che gli esseri umani godono di uno status morale unico e che il rispetto di questo status morale deve essere garantito in tutti gli aspetti della vita. Facendo eco all'imperativo categorico kantiano, la dignità umana è specificata dall'AI HLEG come l'esigenza di trattare gli esseri umani come soggetti, piuttosto che come oggetti (la terminologia kantiana sarebbe quella di trattare le persone come "fini" anziché "mezzi puri"). Sebbene la dignità umana sia attribuita a ciascun essere umano in quanto soggetto morale individuale, l'HLEG non la considera una struttura individualistica. Al contrario, le linee guida interpretano questo specifico approccio antropocentrico come un modo per garantire che i sistemi di intelligenza artificiale siano al servizio "dell'umanità e del bene comune" e considerino gli individui nel loro contesto economico, sociale e politico. Questo approcciolegato alla dignità umana ispira i quattro principi etici che governeranno gli usi dell'IA, vale a dire rispetto per l'autonomia umana, prevenzione dei danni, equità e spiegabilità.

Al terzo livello, le linee guida trasformano i principi etici citati in requisiti chiave per sistemi di IA affidabili (cioè etici e robusti). La logica di questo ulteriore passo è quella di specificare i principi etici al fine di renderli attuabili nella pratica. Dal principio etico del rispetto dell'autonomia umana, quindi,l'AI HLEG deriva il requisito dell'azione umana e della supervisione sugli artefatti dell'IA. Secondo questo requisito, l'intelligenza artificiale non dovrebbe sostituire, ma piuttosto supportare il processo decisionale umano, favorendo l'azione degli utenti. Date le norme e i regolamenti rigidi e ben consolidati delle applicazioni sanitarie, una volta che un prodotto MB raggiunge la fase di sperimentazione clinica, le questioni relative al benessere individuale saranno al centro dell'attenzione, indipendentemente dal fatto che tale principio sia incluso o meno in un elenco di etiche tecnologiche emergenti. Tuttavia, quando lo sviluppo è ancora nelle mani dei biologi piuttosto che dei ricercatori medici, è certamente opportuno prestare attenzione tempestivamente a questi problemi.

4. Questioni etiche legate all'IA

È essenziale considerare le questioni etiche e di privacy quando si utilizza l'IA nelle biotecnologie. Ciò significa garantire che i dati delle persone vengano utilizzati in modo consensuale e anonimo, rispettando i loro diritti. Adottando un approccio etico e responsabile all'uso dell'IA, possiamo sfruttarne il potenziale in modo sicuro e vantaggioso per la società, promuovendo così l'innovazione eil progresso nel campo delle biotecnologie. In questo contesto, la questione della responsabilità emerge come uno dei problemi più complessi e delicati da affrontare, sia a livello di policy che di strategia legislativa. L'idea che il software possa apprendere dai dati e modificare i propri algoritmi per sfruttare ancora di più il suo potenziale a lungo termine è fondamentale nell'apprendimento automatico. Tuttavia, l'apprendimento automatico non ha solo un accesso infinito ai database, ma anche grandi restrizioni e limiti decisivi che difficilmente possono essere superati.

4.1. L’intelligenza artificiale non può “pensare” in modo astratto

L’intelligenza artificiale non è in grado di reagire in modo flessibile a piccoli cambiamenti o a nuovi problemi di applicazione. Non può riconoscere e stabilire alcun collegamento logico o casuale in concetti astratti. Non può trasferire le conoscenze che ha appreso ad altri livelli. Questa incapacità di adattarsi a circostanze cambiate è dovuta al fatto che l’IA “impara” solo ciò che è disposto nella programmazione. Tuttavia, la codifica può fornire conclusioni astratte solo in misura limitata, in modo che la flessibilità del “pensiero” logico non possa essere realizzata. Per questo motivo, ad esempio, il riconoscimento di modelli è la loro migliore abilità, che gli esseri umani non potrebbero mai acquisire nella quantità di set di dati e, soprattutto, nel breve lasso di tempo.

Il funzionamento del cervello umano è estremamente complesso, tanto che può derivare tratti logici e concetti astratti senza dover apprendere ogni problema da zero. Ma l’IA non può mettere in discussione le conclusioni o pensare in modo critico alle cose.

4.2. Problema di ricorsività

Le macchine possono apprendere e diventare “più intelligenti”, ma non hanno la capacità di creare da sole una macchina ancora più potente. Non esiste un’IA che possa migliorare sé stessa. Solo gli esseri umani possono utilizzare le loro abilità cognitive e l’intelligenza creativa e associativa per concepire e costruire macchine ottimizzate. Di conseguenza, l’apprendimento automatico è limitato solo all’aumento della competenza e della velocità di apprendimento. É importante sottolineare che questo rappresenta un limite significativo nell’attuale stato dell’IA. Nonostante la potenza computazionale e l’abilità di elaborare enormi quantità di dati, l’IA non possiede ancora l’autocoscienza o la capacità di auto-migliorarsi senza l’intervento umano. Ciò sottolinea l’importanza del ruolo umano nello sviluppo e nell’ottimizzazione continua dell’intelligenza artificiale.

4.3. Problema di trasparenza delle decisioni della macchina

C’è un grande deficit nella tracciabilità del processo decisionale dell’IA. Come e perché l’IA decida non può essere illustrato, in quanto essa raffigura e seleziona diversi filoni all’interno del processo di apprendimento fino a giungere a un risultato. Questa selezione non può essere visualizzata in nessun momento, complicando così la trasparenza continua che è indispensabile nelle attività competitive. Si sta già lavorando per scomporre il processo di apprendimento dell’IA in singoli strumenti.

Tuttavia, ci vorrà del tempo per raggiungere una svolta in questo senso.

É importante capire che la mancanza di trasparenza nelle decisioni prese dall’IA rappresenta un problema significativo, noto come “scatola nera”. Questo implica che le decisioni prese da un sistema di intelligenza artificiale possono essere complesse e non facilmente comprensibili per gli umani. Inoltre, questa mancanza di trasparenza può portare a problemi etici e di responsabilità, specialmente quando l’IA viene utilizzata in settori critici come la medicina o la sicurezza.

4.4. Limiti morali ed etici

L’IA seleziona solo un singolo elemento di informazione rilevante da tali insiemi di dati, ma non può valutare più contenuti e pensare in modo associativo, come facciamo noi umani. Questo comporta un’elevata probabilità di errore. Ad esempio, l’IA ha difficoltà a riconoscere idiomi o formulazioni discriminatorie. Infatti le conoscenze specifiche per età, genere o religione così come quelle etiche non possono semplicemente essere codificate e rese “riconoscibili” in insiemi di dati. Apprendiamo queste informazioni in modo associativo e

situazionale nel corso della nostra vita. Questa capacità cognitiva è socialmente fondamentale. L'IA non può fare questo. Non sa cosa sia un insulto. Se, ad esempio, l'IA in bot sociali o di gioco riproduce informazioni che sono discriminatorie e denigratorie, ne derivano rischi imprenditoriali e gravi problemi di immagine. Il problema è che l'IA non sa fare di meglio perché non può comprendere questi meccanismi. Essa infatti, come qualsiasi altro strumento tecnologico, non è intrinsecamente buona o cattiva, ma è il suo uso che può portare a risultati positivi o negativi. Questo mette in evidenza l'importanza di un'etica dell'IA solida e ben considerata, che tenga conto non solo della precisione e dell'efficienza, ma anche dei valori umani fondamentali come l'equità, la privacy e la dignità. È necessario un maggiore impegno per garantire che l'IA sia progettata e utilizzata in modo da rispettare questi principi, e per rispondere in modo efficace quando si verificano violazioni.

5. Conclusioni

Esistono diversi temi caldi nel campo dell'intelligenza artificiale e della biotecnologia che sono attualmente oggetto di ricerca attiva e che probabilmente continueranno a essere aree di interesse infuturo:

- AI/ML e analisi dei dati: stanno diventando sempre più importanti nella biotecnologia, poiché possono essere utilizzate per analizzare grandi set di dati e fare previsioni su sistemi biologici complessi. Ciò include, ad esempio, l'uso di tecniche di intelligenza artificiale per analizzare dati genomici.
- Scoperta e sviluppo di farmaci: l'intelligenza artificiale può essere utilizzata per analizzare grandi quantità di dati per identificare modelli e relazioni che potrebbero non essere evidenti agli esseri umani. Ciò può essere utilizzato per identificare nuovi farmaci e bersaglifarmacologici, nonché per ottimizzare le terapie esistenti.
- Medicina personalizzata: l'intelligenza artificiale può essere utilizzata per analizzare i dati genomici di un individuo e altri tipi di dati sanitari per sviluppare piani di trattamento personalizzati adattati alle sue esigenze specifiche. Ciò include l'uso di algoritmi di apprendimento automatico per prevedere la risposta di un individuo a un particolare trattamento e identificare potenziali reazioni avverse.
- Diagnostica e previsione delle malattie: l'intelligenza artificiale può essere utilizzata per analizzare dati provenienti da varie fonti, ad esempio cartelle cliniche elettroniche e dispositivi indossabili, per identificare modelli e correlazioni che potrebbero indicare la presenza di una particolare malattia. Ciò può contribuire a migliorare l'accuratezza delle diagnosi e consentire interventi tempestivi per prevenire la progressione delle malattie.
- Analisi delle immagini biomediche: l'intelligenza artificiale può essere utilizzata per analizzare immagini mediche, come scansioni TC e immagini MRI, per identificare anomalie e diagnosticare malattie. Ciò include l'uso di algoritmi di deep learning per segmentare e classificare automaticamente le strutture nelle immagini mediche.

La digitalizzazione e la trasformazione digitale sono quindi fondamentali per l'inizio di qualsiasi applicazione dell'intelligenza artificiale. I fattori decisivi sono stati e continueranno a essere proprio la disponibilità di volumi di dati grandi e di alta qualità e il rapido aumento della potenza di calcolo. Il nostro desiderio per il futuro è l'equità dell'IA e lo sviluppo di una scienza sempre più libera che possa permettere lo sviluppo di ecosistemi intelligenza artificiale che siano a beneficio del pianeta intero.

6. Bibliografia

1. www.intelligenzaartificiale.it/
2. https://letteredallafacolta.univpm.it/intelligenza-artificiale-scienza-e-tecnologia-etica-e-impatto-nella-sanita/
3. www.medialaws.eu/wp-content/uploads/2022/01/3-21-Portinale.pdf
4. www.toptut.com/it/artificial-intelligence-in-biotech/
5. www.microbiologiaitalia.it/bioinformatica/ia-in-biotecnologia-lintelligenza-artificiale-nelle-scienze-della-vita
6. https://disf.org/files/pt-13_benanti.pdf
7. www.thescienceofwheremagazine.it/2023/10/28/lintegrazione-di-intelligenza-artificiale-e-biotecnologie/
8. https://acrobat.adobe.com/id/urn:aaid:sc:EU:eda910ad-a6ee-4052-9ce4-3603db5fdc5b
9. www.sciencedirect.com/science/article/pii/S266618882200020X
10. http://bioetica.governo.it

ESPLORAZIONI ETICHE NELL'ERA DELL'INTELLIGENZA ARTIFICIALE: RIFLESSIONI E PROSPETTIVE

Andrea Romano V - Dsc



Introduzione all'intelligenza artificiale (IA) e alla sua rapida diffusione nella società contemporanea.

L'intelligenza artificiale (IA) è una delle più potenti e trasformative tecnologie del nostro tempo, caratterizzata dalla capacità di sistemi informatici di emulare l'intelligenza umana per compiere una vasta gamma di compiti. Questa disciplina multidisciplinare si basa su algoritmi, dati e potenza di calcolo per apprendere da esperienze passate, riconoscere schemi, prendere decisioni e adattarsi all'ambiente circostante. La sua rapida diffusione nella società contemporanea ha portato a cambiamenti radicali nei settori economici, sociali e culturali, aprendo nuove opportunità e sollevando importanti questioni etiche e morali.

Le radici concettuali dell'IA risalgono a diverse discipline, tra cui la filosofia, la matematica, la psicologia e la cibernetica. Negli anni '50 del XX secolo, il termine "intelligenza artificiale" fu coniato e divenne oggetto di studio accademico e ricerca. I primi approcci all'IA si concentravano su algoritmi simbolici, come la logica formale e il reasoning basato su regole. Con il passare degli anni, l'IA ha subito una serie di evoluzioni, passando dall'approccio simbolico all'approccio statistico e al machine learning. Quest'ultimo si è

dimostrato particolarmente efficace nell'addestramento di modelli predittivi su grandi quantità di dati, aprendo la strada alla realizzazione di applicazioni sempre più sofisticate, come il riconoscimento vocale, la traduzione automatica, la guida autonoma e molto altro ancora.

L'IA ha trovato applicazioni in una vasta gamma di settori, tra cui l'assistenza sanitaria, la finanza, l'automazione industriale, il marketing, l'istruzione e molti altri. Ad esempio, nell'ambito sanitario, l'IA viene utilizzata per diagnosticare malattie, pianificare trattamenti personalizzati e condurre ricerche mediche avanzate. Nel settore finanziario, gli algoritmi di IA vengono impiegati per analizzare dati finanziari, prevedere tendenze di mercato e mitigare rischi. Tuttavia, l'ampia diffusione dell'IA ha anche generato preoccupazioni e dibattiti riguardanti il suo impatto sulla società. La disoccupazione tecnologica è una delle principali preoccupazioni, con la prospettiva di molti lavori tradizionali resi obsoleti dall'automazione e dall'IA. Inoltre, l'IA solleva interrogativi sull'equità, la privacy, la sicurezza e il controllo dell'informazione, alimentando un dibattito pubblico su come regolare e gestire questa tecnologia in modo responsabile.

Guardando al futuro, l'IA promette di rivoluzionare ulteriormente la società, con nuove innovazioni che potrebbero trasformare radicalmente il modo in cui lavoriamo, viviamo e interagiamo. Tuttavia, affrontare le sfide etiche associate all'IA diventa sempre più cruciale. È fondamentale garantire che questa tecnologia sia sviluppata e utilizzata in modo etico e responsabile, con un'attenzione particolare all'equità, alla trasparenza, alla responsabilità e alla sicurezza. In conclusione, l'IA rappresenta una delle forze trainanti del cambiamento nella società contemporanea, con il potenziale di portare benefici significativi, ma anche rischi e sfide complesse. Affrontare queste sfide richiederà un approccio collaborativo e multidisciplinare, che coinvolga esperti tecnici, decisori politici, professionisti etici e il pubblico in generale. Solo attraverso un impegno condiviso per un'IA etica e responsabile possiamo garantire che questa tecnologia abbia un impatto positivo e sostenibile sulla nostra società.

Fondamenti dell'etica applicata all'IA: definizioni e concetti fondamentali.

L'intelligenza artificiale (IA) è una tecnologia in continua evoluzione che solleva una serie di questioni etiche e morali nel suo sviluppo, implementazione e utilizzo. Affrontare tali questioni richiede una comprensione approfondita dei fondamenti dell'etica applicata all'IA, così come una riflessione critica sui principi e i valori che dovrebbero guidare le nostre azioni e decisioni nel contesto dell'IA.

Per comprendere appieno l'etica applicata all'IA, è fondamentale iniziare con una chiara definizione di termini chiave. L'etica stessa è una branca della filosofia che si occupa dello studio dei principi morali che guidano il comportamento umano e delle questioni relative al giusto e all'ingiusto, al bene e al male. Essa fornisce il quadro concettuale per valutare le azioni umane in relazione ai loro effetti sull'individuo e sulla società nel suo complesso.

L'IA, d'altra parte, si riferisce alla capacità dei sistemi informatici di emulare l'intelligenza umana per compiere una vasta gamma di compiti. Questi sistemi si basano su algoritmi, dati e potenza di calcolo per raggiungere i loro obiettivi, che possono includere il riconoscimento di pattern, l'apprendimento automatico e la presa di decisioni. Tuttavia, l'IA solleva una serie di domande etiche riguardo al modo in cui viene sviluppata, implementata e utilizzata.

Tra i concetti fondamentali dell'etica applicata all'IA vi sono la trasparenza e l'accountability. Questi concetti sottolineano l'importanza di rendere trasparenti i processi decisionali dell'IA e di attribuire responsabilità agli individui o alle organizzazioni per le conseguenze delle azioni dell'IA. La mancanza di trasparenza nei processi decisionali dell'IA può portare a decisioni discriminatorie o dannose, mentre l'assenza di responsabilità può rendere difficile l'individuazione dei responsabili e la correzione degli errori.

Inoltre, l'etica applicata all'IA si occupa di questioni di giustizia e equità. È importante che l'IA non favorisca alcun gruppo sociale o individuale a discapito degli altri. Dovrebbe essere progettata per promuovere la giustizia sociale e l'equità, evitando discriminazioni e pregiudizi. Allo stesso modo, l'IA non dovrebbe limitare l'autonomia umana o influenzare indebitamente le scelte individuali. Gli individui dovrebbero mantenere il controllo sulle decisioni che li riguardano, nonostante l'impiego dell'IA.

La privacy e la sicurezza sono anche concetti fondamentali nell'etica applicata all'IA. L'IA dovrebbe rispettare la privacy e la sicurezza dei dati personali degli individui e deve essere progettata e utilizzata in modo da garantire la protezione delle informazioni sensibili e prevenire abusi o violazioni della privacy. Inoltre, l'IA dovrebbe essere sviluppata e utilizzata per massimizzare i benefici per l'individuo e la società nel suo complesso, evitando danni o effetti negativi indesiderati.

In conclusione, comprendere i fondamenti dell'etica applicata all'IA è essenziale per affrontare le sfide etiche che questa tecnologia solleva nella società contemporanea. I concetti di trasparenza, accountability, giustizia, equità, autonomia, controllo, privacy, sicurezza, beneficenza e non-maleficenza forniscono un quadro concettuale per guidare lo sviluppo, l'implementazione e l'utilizzo responsabile dell'IA. Affrontare queste sfide richiederà un impegno continuo da parte di ricercatori, professionisti, decisori politici e la società nel suo complesso per garantire che l'IA contribuisca al benessere umano e alla giustizia sociale.

L'impatto dell'IA sulla società e sull'occupazione: implicazioni etiche e sociali.

L'intelligenza artificiale (IA) è diventata una forza trainante nella trasformazione della società e dell'economia contemporanea. Questa tecnologia ha dimostrato un potenziale rivoluzionario in molteplici settori, dall'assistenza sanitaria alla finanza, dalla produzione all'educazione. Tuttavia, la diffusione sempre più ampia dell'IA solleva una serie di questioni etiche e sociali, tra cui l'impatto sull'occupazione e sulla forza lavoro.

Impatto sull'Occupazione

Una delle principali preoccupazioni riguarda l'effetto dell'IA sull'occupazione e sulla forza lavoro. L'IA ha la capacità di automatizzare una vasta gamma di compiti e processi, il che potrebbe portare alla riduzione della domanda di lavoro in settori che coinvolgono attività ripetitive e standardizzate. Questo fenomeno, noto come disoccupazione tecnologica, solleva interrogativi sulla sicurezza economica e sulle prospettive future dei lavoratori coinvolti.

Tuttavia, è importante sottolineare che l'IA non è solo distruttiva, ma può anche creare nuove opportunità occupazionali. L'implementazione dell'IA può generare nuove professioni legate alla progettazione, sviluppo, manutenzione e gestione dei sistemi AI. Ad esempio, sono richiesti specialisti in machine learning, ingegneri dell'IA, esperti di etica dell'IA e molti altri. Inoltre, l'IA può migliorare la produttività e la competitività delle imprese, creando nuove opportunità di lavoro in settori adiacenti o emergenti.

Per mitigare l'impatto negativo sull'occupazione, sono necessarie politiche e misure che favoriscano la riqualificazione e la formazione professionale continua. Investire in programmi di educazione e addestramento che sviluppino competenze adatte all'economia digitale del futuro può aiutare i lavoratori ad adattarsi alle nuove sfide e opportunità presentate dall'IA. Inoltre, è importante garantire un equo accesso alle opportunità di formazione e all'occupazione nell'ambito dell'IA, al fine di ridurre le disuguaglianze sociali ed economiche.

Impatto Sociale ed Etico

Al di là delle implicazioni sull'occupazione, l'IA ha un impatto più ampio sulla società nel suo complesso. Ad esempio, l'IA può influenzare la distribuzione delle risorse e la decisione politica, attraverso l'automazione dei processi decisionali e l'analisi dei dati. Ciò solleva interrogativi riguardo alla responsabilità, alla trasparenza e alla giustizia nell'allocazione delle risorse e delle opportunità.

Inoltre, l'IA può contribuire alla creazione o all'accentuazione di disuguaglianze sociali ed economiche. Se non regolamentata adeguatamente, l'IA potrebbe amplificare i pregiudizi e le discriminazioni presenti nei dati utilizzati per l'addestramento dei modelli AI, portando a decisioni ingiuste o discriminatorie. Ciò solleva la necessità di sviluppare e applicare algoritmi etici e equi che tengano conto delle disparità e promuovano la giustizia sociale.

In conclusione, l'IA ha un impatto profondo e pervasivo sulla società e sull'occupazione, con implicazioni etiche e sociali significative. Affrontare queste sfide richiede un approccio olistico e collaborativo che coinvolga governi, istituzioni, aziende e la società nel suo complesso. È cruciale adottare politiche e pratiche che promuovano un utilizzo responsabile dell'IA, che massimizzi i suoi benefici e minimizzi i rischi per il benessere umano e per la coesione sociale.

Responsabilità nell'IA: chi è responsabile per le azioni e le decisioni dell'IA?

L'Intelligenza Artificiale (IA) rappresenta una delle tecnologie più avanzate e promettenti del nostro tempo, con un potenziale enorme per migliorare la nostra vita quotidiana, trasformare industrie intere e risolvere problemi complessi. Tuttavia, insieme a questo potenziale, sorgono anche questioni etiche e legali che richiedono un'analisi approfondita e una risposta adeguata. Una di queste questioni fondamentali riguarda la responsabilità per le azioni e le decisioni dell'IA.

Complessità della Responsabilità nell'IA

La natura stessa dell'IA, con la sua capacità di apprendere dai dati e di prendere decisioni autonomamente, rende la questione della responsabilità estremamente complessa. A differenza delle tecnologie tradizionali, dove è più semplice individuare i soggetti responsabili, nell'IA si aprono una serie di sfide nel determinare chi debba essere ritenuto responsabile in caso di danni o conseguenze negative.

Ad esempio, consideriamo un sistema di IA utilizzato per prendere decisioni di assunzione in un'azienda. Se questo sistema adotta criteri discriminatori basati su dati storici sbagliati, chi è responsabile? Gli sviluppatori dell'algoritmo, i proprietari dei dati utilizzati per addestrare il modello, o l'azienda stessa che ha implementato il sistema? Questo è solo uno dei numerosi scenari complessi che sollevano interrogativi sulla responsabilità nell'IA.

Responsabilità Legale e Normativa

Dal punto di vista legale, stabilire la responsabilità nell'IA richiede un aggiornamento e un adattamento delle normative esistenti per tener conto delle peculiarità di questa tecnologia. Le leggi devono chiarire chi sia considerato responsabile in diversi contesti, definendo le linee guida per la gestione dei rischi e la compensazione delle vittime di danni causati dall'IA.

In questo contesto, emergono concetti come la responsabilità oggettiva, che potrebbe attribuire la responsabilità a chiunque abbia il controllo o la supervisione diretta del sistema di IA, e la responsabilità condivisa, che coinvolge tutti gli attori coinvolti nel ciclo di vita dell'IA, compresi gli sviluppatori, i fornitori di dati, gli utenti finali e i regolatori.

Tuttavia, anche se è fondamentale stabilire una cornice normativa chiara, è altrettanto importante riconoscere che la legge da sola potrebbe non essere sufficiente a rispondere completamente alle complessità della responsabilità nell'IA. È necessario sviluppare anche strumenti etici e pratiche aziendali che promuovano la responsabilità e l'accountability nell'implementazione e nell'utilizzo dell'IA.

Implicazioni Etiche e Sociali

Oltre alle considerazioni legali, la questione della responsabilità nell'IA solleva anche importanti questioni etiche e sociali. Ad esempio, chi dovrebbe assumersi la responsabilità morale per le conseguenze delle decisioni dell'IA? Gli sviluppatori che progettano gli algoritmi, i proprietari dei dati utilizzati per addestrare i modelli di IA, o gli utilizzatori finali che prendono decisioni basate sulle raccomandazioni dell'IA?

Inoltre, la responsabilità nell'IA può influenzare la fiducia del pubblico nell'utilizzo di questa tecnologia. Se gli individui percepiscono che non esiste una chiara responsabilità per le azioni dell'IA, potrebbero essere riluttanti ad accettare e ad adottare questa tecnologia, limitando così il suo potenziale benefico per la società.

Strategie per Affrontare la Responsabilità nell'IA

Affrontare la questione della responsabilità nell'IA richiede un approccio olistico e collaborativo che coinvolga tutti gli attori coinvolti, compresi sviluppatori, utenti finali, regolatori e la società nel suo complesso. È essenziale sviluppare normative chiare e contestualmente sensibili che definiscano chi sia responsabile in diverse situazioni e stabiliscano meccanismi di rendicontazione e risarcimento adeguati.

Inoltre, è importante promuovere la trasparenza e l'accountability nell'IA, consentendo una maggiore comprensione e controllo delle decisioni prese da questi sistemi. Ciò può includere la documentazione e la divulgazione delle metodologie utilizzate per addestrare gli algoritmi, nonché la creazione di meccanismi di supervisione e controllo che consentano agli utenti di intervenire o correggere le decisioni dell'IA quando necessario.

In conclusione, la questione della responsabilità nell'IA è una delle sfide più urgenti e complesse che la società contemporanea deve affrontare. Per garantire un utilizzo responsabile e sostenibile dell'IA, è necessario sviluppare normative, meccanismi e pratiche che promuovano la trasparenza, l'accountability e la fiducia del pubblico nell'utilizzo di questa tecnologia emergente. Solo attraverso un impegno condiviso e collaborativo possiamo garantire che l'IA abbia un impatto positivo e duraturo sulla nostra società.

Trasparenza e accountability nell'IA: accesso ai dati, algoritmi e processi decisionali

L'Intelligenza Artificiale (IA) si sta diffondendo sempre più nella nostra società, influenzando molteplici aspetti della nostra vita quotidiana, dal modo in cui interagiamo con i dispositivi digitali alla gestione dei servizi pubblici. Tuttavia, con questa crescente diffusione sorgono anche preoccupazioni etiche e sociali riguardanti la trasparenza e l'accountability nell'IA. In questo capitolo, esploreremo approfonditamente l'importanza dell'accesso ai dati, alla trasparenza degli algoritmi e alla responsabilità nei processi decisionali dell'IA.

Comprensione dell'Importanza della Trasparenza e dell'Accountability nell'IA

Per comprendere appieno l'importanza della trasparenza e dell'accountability nell'IA, è fondamentale definire chiaramente questi concetti. La trasparenza si riferisce alla chiarezza e all'accessibilità delle informazioni riguardanti il funzionamento degli algoritmi di IA, dei dati utilizzati per addestrarli e dei processi decisionali derivati da tali algoritmi. L'accountability,

d'altra parte, riguarda l'attribuzione di responsabilità per le azioni e le decisioni prese dall'IA, garantendo che gli attori coinvolti siano tenuti responsabili delle conseguenze delle loro azioni.

Accesso ai Dati nell'IA

Uno degli aspetti cruciali della trasparenza nell'IA è l'accesso ai dati utilizzati per addestrare gli algoritmi. È fondamentale che gli sviluppatori e gli utenti finali possano accedere ai dati utilizzati per addestrare gli algoritmi di IA al fine di valutarne l'accuratezza, la rappresentatività e la potenziale presenza di bias. Senza accesso a questi dati, diventa difficile comprendere e valutare le decisioni prese dall'IA e garantire la loro equità e giustizia.

Trasparenza degli Algoritmi nell'IA

Oltre all'accesso ai dati, è importante anche garantire la trasparenza degli algoritmi stessi. Questo significa che gli sviluppatori devono fornire informazioni dettagliate sul funzionamento degli algoritmi di IA, inclusi i criteri utilizzati per prendere decisioni e le caratteristiche dei dati che influenzano tali decisioni. La trasparenza degli algoritmi consente una maggiore comprensione delle decisioni dell'IA da parte degli utenti finali e può contribuire a individuare e correggere eventuali errori o bias presenti nei sistemi di IA.

Accountability nei Processi Decisionali dell'IA

Infine, la responsabilità nei processi decisionali dell'IA è un altro aspetto critico della trasparenza e dell'accountability nell'IA. È importante stabilire chi sia responsabile delle decisioni prese dall'IA e garantire che gli attori coinvolti siano tenuti responsabili delle conseguenze delle loro azioni. Questo può includere gli sviluppatori degli algoritmi, i fornitori di dati, gli utenti finali e le organizzazioni che implementano e utilizzano l'IA nei loro processi decisionali.

Promuovere la Trasparenza e l'Accountability nell'IA

Per promuovere la trasparenza e l'accountability nell'IA, è necessario adottare una serie di misure e pratiche. Queste possono includere la creazione di standard e linee guida per la documentazione e la divulgazione dei dati e degli algoritmi utilizzati nell'IA, l'implementazione di meccanismi di supervisione e controllo per monitorare le decisioni dell'IA e la promozione di una cultura di responsabilità e trasparenza all'interno delle organizzazioni che sviluppano e utilizzano l'IA.

In conclusione, la trasparenza e l'accountability sono elementi essenziali per garantire un utilizzo responsabile e sostenibile dell'IA. Garantire l'accesso ai dati, la trasparenza degli algoritmi e la responsabilità nei processi decisionali sono passi fondamentali per promuovere la fiducia del pubblico e assicurare che l'IA sia utilizzata in modo equo, giusto ed etico. Solo attraverso un impegno condiviso e collaborativo possiamo garantire che l'IA abbia un impatto positivo e duraturo sulla nostra società.

Equità e giustizia nell'IA: come evitare o mitigare i pregiudizi e le discriminazioni algoritmiche.

L'Intelligenza Artificiale (IA) rappresenta una delle tecnologie più potenti e transformative del nostro tempo, promettendo di rivoluzionare praticamente ogni settore della nostra società. Tuttavia, insieme alle promesse di efficienza e innovazione, l'IA porta con sé delle profonde preoccupazioni riguardanti l'equità e la giustizia nel processo decisionale

algoritmico. In un mondo in cui l'IA influenza le nostre vite in modi sempre più profondi, è cruciale comprendere e affrontare le questioni relative ai pregiudizi e alle discriminazioni algoritmiche che potrebbero minacciare la nostra società inclusiva e giusta.

Per affrontare la questione dell'equità e della giustizia nell'IA, dobbiamo innanzitutto comprendere le radici dei pregiudizi e delle discriminazioni algoritmiche. Queste problematiche emergono spesso a causa della presenza di bias nei dati utilizzati per addestrare gli algoritmi di IA o nelle regole di decisione implementate all'interno degli stessi algoritmi. Tali bias possono derivare da una varietà di fonti, tra cui dati storici discriminatori, mancanza di rappresentatività dei dati o persino bias impliciti degli sviluppatori.

Le discriminazioni algoritmiche rappresentano una minaccia significativa per la nostra società. Esse possono perpetuare le disuguaglianze esistenti, amplificare le disparità sociali ed economiche e privare gruppi svantaggiati di opportunità e risorse. Ad esempio, un algoritmo di selezione del personale che favorisce inconsciamente i candidati di una determinata etnia, genere o classe sociale potrebbe escludere ingiustamente individui qualificati e contribuire all'ingiustizia nell'accesso al lavoro.

Per contrastare le discriminazioni algoritmiche, è necessario adottare una serie di strategie preventive e correttive. Innanzitutto, è fondamentale garantire la diversità e la rappresentatività dei dati utilizzati per addestrare gli algoritmi di IA. Questo può essere ottenuto attraverso la raccolta di dati diversificati e il controllo attivo dei bias durante il processo di addestramento.

In secondo luogo, è importante sviluppare e applicare algoritmi etici che tengano conto delle disparità sociali ed economiche esistenti. Questo può includere l'implementazione di tecniche di fair machine learning che mirano a garantire equità e giustizia nei risultati prodotti dagli algoritmi, riducendo al minimo l'impatto dei bias nei dati.

Inoltre, è cruciale promuovere la trasparenza e la responsabilità nell'utilizzo dell'IA. Gli sviluppatori devono essere trasparenti riguardo ai criteri utilizzati dagli algoritmi e devono essere pronti a rendere conto delle conseguenze delle loro decisioni. L'implementazione di meccanismi di supervisione e controllo può aiutare a monitorare e correggere eventuali discriminazioni algoritmiche nel tempo.

Infine, è essenziale coinvolgere le parti interessate, inclusi gli esperti di etica, i membri della comunità e gli utenti finali, nel processo decisionale riguardante l'utilizzo dell'IA. Questo può contribuire a garantire che le decisioni siano prese in modo consapevole e inclusivo, tenendo conto delle diverse prospettive e esigenze delle persone coinvolte.

In conclusione, l'equità e la giustizia nell'IA sono fondamentali per garantire che questa tecnologia contribuisca a promuovere una società più equa e inclusiva. Evitare o mitigare i pregiudizi e le discriminazioni algoritmiche richiede un impegno condiviso e collaborativo da parte di sviluppatori, regolatori, organizzazioni e la società nel suo complesso. Solo attraverso un approccio olistico e responsabile possiamo garantire che l'IA sia utilizzata per il bene comune, evitando gli effetti negativi delle discriminazioni algoritmiche e contribuendo a costruire un mondo più equo e giusto per tutti.

Privacy e sicurezza nell'IA: protezione dei dati personali e rischi di violazioni della privacy.

L'Intelligenza Artificiale (IA) è diventata un pilastro fondamentale nella nostra società moderna, permeando molteplici settori con la sua presenza onnipresente. Tuttavia, con la sua diffusione, emergono crescenti preoccupazioni riguardo alla privacy e alla sicurezza dei dati personali. La capacità dell'IA di analizzare enormi quantità di dati, inclusi dati altamente sensibili, solleva interrogativi cruciali sulla protezione dei dati e sul rischio di violazioni della privacy.

La protezione dei dati personali è di vitale importanza nell'ambito dell'IA. Dati sensibili come informazioni personali, dati medici o finanziari richiedono un trattamento

particolarmente attento per garantire che rimangano al sicuro da accessi non autorizzati o da utilizzi impropri. Senza adeguati protocolli di sicurezza, i dati personali possono essere esposti a rischi quali violazioni della privacy, furto di identità o discriminazione.

L'archiviazione e la trasmissione sicura dei dati sono essenziali per prevenire violazioni della privacy nell'ambito dell'IA. L'uso diffuso della crittografia può garantire che i dati siano inaccessibili a individui non autorizzati, proteggendo così la privacy degli utenti. Inoltre, protocolli robusti per la gestione degli accessi e il controllo degli utenti possono limitare l'accesso ai dati solo a coloro che ne hanno il diritto.

Nonostante gli sforzi per proteggere i dati personali, rimangono numerosi rischi di violazioni della privacy nell'ambito dell'IA. La raccolta e l'elaborazione massiccia di dati aumentano il rischio di accessi non autorizzati o di fughe di dati. Gli algoritmi di IA stessi possono essere vulnerabili a manipolazioni malevole o attacchi informatici, compromettendo così la sicurezza dei dati e la privacy degli utenti.

Affrontare questi rischi richiede una combinazione di misure preventive e correttive. Le aziende e le organizzazioni devono investire in tecnologie avanzate di sicurezza informatica, come sistemi di rilevamento delle intrusioni e firewall avanzati, per proteggere i sistemi di IA da potenziali minacce. Allo stesso tempo, è essenziale promuovere una cultura della privacy e della sicurezza all'interno delle organizzazioni, garantendo che il personale sia adeguatamente informato sui rischi associati alla gestione dei dati personali e sulle best practice per mitigare tali rischi.

Coinvolgere attivamente gli utenti è un altro aspetto cruciale per migliorare la privacy e la sicurezza nell'IA. Gli utenti devono essere informati sui dati che vengono raccolti, su come vengono utilizzati e su come possono esercitare il controllo sulla propria privacy. L'adozione di politiche trasparenti e di procedure di sicurezza, insieme a meccanismi di consenso esplicito, può contribuire a costruire la fiducia del pubblico nell'utilizzo dell'IA.

In conclusione, la privacy e la sicurezza dei dati personali sono questioni di cruciale importanza nell'ambito dell'IA. Proteggere i dati personali degli utenti e prevenire le violazioni della privacy sono imperativi etici e legali che richiedono un impegno continuo da parte di aziende, governi e individui. Solo attraverso una collaborazione globale e un impegno condiviso possiamo garantire un utilizzo responsabile e sostenibile dell'IA, preservando nel contempo i diritti e le libertà individuali nella società digitale del futuro. Autonomia e controllo nell'IA: bilanciare l'autonomia dell'IA con la supervisione umana.

L'Intelligenza Artificiale (IA) sta progressivamente acquisendo un grado sempre maggiore di autonomia e capacità decisionale, sollevando interrogativi cruciali riguardo al bilanciamento dell'autonomia dell'IA con la supervisione umana. Se da un lato l'autonomia dell'IA può portare a efficienza e innovazione, dall'altro solleva preoccupazioni riguardo alla responsabilità, alla trasparenza e alla sicurezza dei risultati prodotti. In questo capitolo, esamineremo in dettaglio le sfide e le opportunità legate all'autonomia e al controllo nell'IA, esplorando strategie per garantire un equilibrio appropriato tra l'autonomia delle macchine e il ruolo della supervisione umana.

L'autonomia dell'IA rappresenta un'evoluzione significativa nella capacità delle macchine di prendere decisioni senza un intervento umano diretto. Questa autonomia può manifestarsi in diversi contesti, dall'automazione dei processi decisionali nelle industrie alla guida autonoma dei veicoli. Tuttavia, una maggiore autonomia comporta anche una maggiore responsabilità, poiché le decisioni prese dall'IA possono avere conseguenze significative sulla vita e sul benessere delle persone.

La supervisione umana rimane essenziale per garantire un utilizzo responsabile e sicuro dell'IA. Gli esseri umani sono in grado di fornire contesto, giudizio etico e comprensione del contesto sociale che possono essere al di là delle capacità dell'IA. La supervisione umana è particolarmente importante in situazioni complesse o eticamente sensibili, dove è necessaria una valutazione umana per prendere decisioni informate.

Bilanciare l'autonomia dell'IA con la supervisione umana richiede l'adozione di approcci

multifaceted e multidisciplinari. Innanzitutto, è necessario sviluppare e applicare standard etici e normative legali per guidare l'implementazione dell'IA in modo responsabile. Questi standard dovrebbero includere linee guida per la supervisione umana, la trasparenza degli algoritmi e la responsabilità nell'uso dell'IA. In secondo luogo, è essenziale integrare meccanismi di controllo e monitoraggio che consentano agli esseri umani di intervenire quando necessario per correggere eventuali errori o bias nell'IA. Questi meccanismi possono includere sistemi di allerta tempestiva, procedure di revisione umana e interfacce utente intuitive che consentano agli utenti di comprendere e influenzare il comportamento dell'IA.

Inoltre, è importante promuovere la formazione e la consapevolezza tra gli utenti e gli sviluppatori sull'importanza della supervisione umana nell'IA. Questo può essere realizzato attraverso programmi educativi, corsi di formazione e iniziative di sensibilizzazione che enfatizzano il ruolo critico che gli esseri umani devono svolgere nel garantire l'utilizzo etico e sicuro dell'IA.

Infine, è fondamentale incoraggiare la collaborazione e il dialogo tra diverse parti interessate, compresi sviluppatori, regolatori, esperti di etica e rappresentanti della società civile. Solo attraverso un approccio collaborativo e inclusivo possiamo affrontare efficacemente le sfide e le opportunità legate all'autonomia e al controllo nell'IA, garantendo nel contempo che questa tecnologia sia utilizzata per il bene comune e nel rispetto dei valori umani fondamentali.

In conclusione, bilanciare l'autonomia dell'IA con la supervisione umana è una sfida complessa che richiede un approccio olistico e collaborativo. Mentre l'autonomia dell'IA offre opportunità di innovazione e progresso, è fondamentale garantire che questa autonomia sia guidata da principi etici, normative legali e supervisione umana. Solo attraverso un impegno congiunto da parte di tutte le parti interessate possiamo garantire che l'IA sia utilizzata in modo responsabile e sostenibile, contribuendo al benessere e al progresso della società nel suo complesso.

Impatti psicologici e culturali dell'IA: riflessioni sull'interazione umano-macchina.

L'Intelligenza Artificiale (IA), con la sua crescente pervasività nella nostra vita quotidiana, ha portato significativi cambiamenti non solo a livello tecnologico ed economico, ma anche sul piano psicologico e culturale. L'interazione sempre più frequente con sistemi di IA solleva una serie di questioni profonde riguardanti il modo in cui gli esseri umani percepiscono sé stessi, gli altri e il mondo che li circonda.

L'interazione con l'IA può avere un impatto notevole sul benessere psicologico degli individui. Da un lato, l'IA offre possibilità di semplificare compiti e di migliorare l'efficienza, portando a una sensazione di soddisfazione e realizzazione personale. Tuttavia, dall'altro lato, la crescente automazione e l'avanzamento dell'IA possono generare ansia e insicurezza riguardo al futuro del lavoro e al ruolo dell'umanità nel mondo tecnologicamente avanzato.

In particolare, la paura della sostituzione da parte delle macchine può portare a sentimenti di inadeguatezza e perdita di senso di scopo. Gli individui possono trovarsi a lottare con domande esistenziali riguardo al proprio valore e alla propria identità in un mondo in cui le macchine assumono sempre più compiti precedentemente riservati agli esseri umani.

A livello culturale, l'IA sta trasformando le dinamiche sociali e il modo in cui interagiamo e consumiamo informazioni e contenuti digitali. Gli algoritmi di raccomandazione personalizzata e di filtraggio delle informazioni possono creare bolle informative, limitando l'esposizione delle persone a punti di vista diversi e alimentando la polarizzazione delle opinioni.

Inoltre, l'IA sta influenzando la produzione culturale e artistica. Gli algoritmi di

generazione di contenuti possono essere utilizzati per creare opere d'arte, musica e letteratura, sollevando domande riguardo alla natura stessa della creatività e all'autenticità delle opere prodotte dall'IA.

Questo solleva interrogativi profondi sul ruolo degli esseri umani come creatori e sulla perdita di senso di unicità nell'espressione artistica.

Per affrontare gli impatti psicologici e culturali dell'IA, è necessario adottare un approccio equilibrato che tenga conto delle sfide e delle opportunità che questa tecnologia offre. Ciò include l'implementazione di politiche e normative che proteggano i diritti e il benessere degli individui, nonché lo sviluppo di programmi educativi e di sensibilizzazione che promuovano una comprensione consapevole e critica dell'IA e dei suoi impatti sulla società.

In conclusione, gli impatti psicologici e culturali dell'IA sollevano una serie di questioni complesse che richiedono un'attenzione e una riflessione approfondite. Mentre l'IA offre opportunità di innovazione e progresso, è fondamentale affrontare in modo responsabile e consapevole le sfide che essa presenta per la nostra psicologia individuale e per la nostra cultura collettiva. Solo attraverso un dialogo aperto e inclusivo possiamo garantire che l'IA sia utilizzata per il bene comune, preservando nel contempo i valori umani fondamentali e il tessuto stesso della nostra società.

Prospettive future dell'etica dell'IA: sfide emergenti e possibili sviluppi nel campo dell'IA etica.

L'etica dell'Intelligenza Artificiale (IA) è diventata un argomento di discussione sempre più rilevante, poiché l'IA continua a evolversi e ad essere integrata in una vasta gamma di settori e applicazioni. Guardando al futuro, emergono una serie di sfide etiche e possibilità di sviluppo nel campo dell'IA etica che richiedono un'analisi approfondita e una considerazione attenta.

Sfide Emergenti

1. Bias e Discriminazione: Un'enorme sfida nell'IA è rappresentata dal bias incorporato nei dati e negli algoritmi, che può portare a discriminazioni ingiuste e disuguaglianze. Affrontare questo problema richiede non solo la raccolta di dati più diversificati, ma anche lo sviluppo di algoritmi che siano consapevoli dei bias e capaci di mitigarli.
2. Responsabilità e Accountability: Determinare chi è responsabile per le azioni e le decisioni dell'IA è un'altra sfida cruciale. Con l'autonomia crescente dell'IA, diventa sempre più difficile attribuire la responsabilità per eventuali danni o errori. Sviluppare meccanismi chiari di accountability diventa pertanto essenziale per garantire un utilizzo responsabile dell'IA.
3. Trasparenza e Interpretabilità: Gli algoritmi di IA spesso operano come "scatole nere", il che significa che è difficile comprendere il motivo per cui prendono determinate decisioni. Migliorare la trasparenza e l'interpretabilità degli algoritmi è fondamentale per garantire che le decisioni dell'IA siano comprese e accettabili per gli esseri umani.
4. Privacy e Sicurezza: Con la crescente raccolta e l'elaborazione di enormi quantità di dati personali, è fondamentale proteggere la privacy degli individui e garantire la sicurezza dei loro dati. Sviluppare meccanismi robusti di protezione dei dati diventa cruciale per mitigare i rischi di violazioni della privacy e di abusi dei dati.

Possibili Sviluppi

1. Standard Etici e Normative: Il futuro dell'IA etica potrebbe vedere lo sviluppo e l'implementazione di standard etici e normative globali che guidino l'utilizzo responsabile e sostenibile dell'IA. Questi standard potrebbero essere sviluppati da organizzazioni internazionali, governi e stakeholder del settore.
2. Algoritmi Etici e Fairness by Design: Ci potrebbe essere un movimento verso lo

sviluppo di algoritmi etici che incorporano principi etici fin dall'inizio del processo di progettazione. Questi algoritmi potrebbero essere progettati per massimizzare il benessere sociale e minimizzare il rischio di bias e discriminazione.

3. Partecipazione Pubblica e Coinvolgimento degli Stakeholder: Un maggiore coinvolgimento del pubblico e degli stakeholder nel processo decisionale sull'IA etica potrebbe portare a soluzioni più democratiche e socialmente accettabili. Questo potrebbe includere consultazioni pubbliche, tavoli rotondi e forum di discussione che coinvolgono una vasta gamma di prospettive e punti di vista.

4. Educazione e Formazione: Investire nell'educazione e nella formazione sull'IA etica potrebbe essere cruciale per garantire che gli sviluppatori, gli operatori e gli utenti comprendano appieno le implicazioni etiche dell'IA e siano in grado di prendere decisioni informate e consapevoli.

In conclusione, le prospettive future dell'etica dell'IA sono complesse e stimolanti. Mentre emergono nuove sfide etiche con l'evoluzione dell'IA, ci sono anche opportunità senza precedenti per sviluppare soluzioni innovative e responsabili che guidino l'utilizzo dell'IA per il bene comune. Affrontare queste sfide e sfruttare queste opportunità richiederà un impegno continuo e collaborativo da parte di governi, organizzazioni, sviluppatori e società nel suo complesso.

Conclusioni e raccomandazioni: considerazioni finali sull'importanza dell'etica nell'era dell'IA e suggerimenti per affrontare le sfide etiche in modo efficace.

Nel corso di questo saggio, abbiamo esplorato in profondità le complesse intersezioni tra l'Intelligenza Artificiale (IA) e l'etica, riflettendo sulle sfide e sulle opportunità che questa tecnologia porta con sé. L'IA ha dimostrato un enorme potenziale nel trasformare molteplici aspetti della nostra vita, ma allo stesso tempo ha sollevato una serie di interrogativi fondamentali riguardo alla responsabilità, alla giustizia, alla privacy e alla dignità umana. Navigare in questo terreno etico complesso richiede un approccio olistico e consapevole, che tenga conto dei molteplici interessi in gioco e delle implicazioni a lungo termine delle nostre decisioni.

L'importanza dell'etica nell'era dell'IA non può essere sottovalutata. Come abbiamo visto, l'IA può amplificare e perpetuare pregiudizi e discriminazioni, a meno che non sia guidata da principi etici che pongono al centro il rispetto dei diritti umani e della dignità. Inoltre, l'IA può sollevare interrogativi profondi riguardo alla responsabilità e alla trasparenza delle decisioni, poiché le macchine diventano sempre più autonome e capaci di prendere decisioni che hanno un impatto diretto sulle vite delle persone.

Per affrontare queste sfide in modo efficace, è necessario adottare un approccio proattivo e collaborativo che coinvolga una vasta gamma di stakeholder, tra cui governi, industria, società civile e comunità accademiche. Una delle raccomandazioni chiave è quella di sviluppare e implementare standard etici e normative che guidino lo sviluppo e l'utilizzo responsabile dell'IA. Questi standard dovrebbero includere principi fondamentali come equità, trasparenza, responsabilità e rispetto della privacy.

Inoltre, è essenziale investire nella ricerca e nello sviluppo di strumenti e metodologie che consentano di identificare e mitigare i bias negli algoritmi, così come di garantire la trasparenza e l'interpretabilità delle decisioni dell'IA. Ciò potrebbe includere l'adozione di pratiche come l'Audit dell'IA, che mira a valutare criticamente gli algoritmi per individuare e correggere eventuali pregiudizi o errori.

Un'altra raccomandazione importante è quella di promuovere l'educazione e la consapevolezza pubblica sull'IA e sui suoi impatti etici. La comprensione dei cittadini sul funzionamento dell'IA e sui suoi potenziali impatti può contribuire a informare il dibattito pubblico e a influenzare le decisioni politiche e aziendali in modo più responsabile.

Infine, è cruciale incoraggiare la collaborazione internazionale e la condivisione delle

migliori pratiche nell'ambito dell'IA etica. Poiché l'IA non conosce confini nazionali, è importante lavorare insieme per sviluppare standard e politiche che siano coerenti e armonizzati a livello globale, garantendo così che l'IA sia utilizzata per promuovere il bene comune e il rispetto dei diritti umani in tutto il mondo.

In conclusione, l'era dell'IA offre enormi opportunità di innovazione e progresso, ma al tempo stesso solleva sfide etiche complesse che richiedono una risposta urgente e impegnativa. Affrontare queste sfide richiede un impegno collettivo per garantire che l'IA sia sviluppata e utilizzata in modo responsabile e sostenibile, nel rispetto dei valori umani fondamentali e del benessere della società nel suo complesso. Solo attraverso un approccio etico e consapevole possiamo realizzare pienamente il potenziale trasformativo dell'IA per il bene di tutti.

ROBOTICA EDUCATIVA E SUE APPLICAZIONI NELLA SCUOLA SUPERIORE. ESPERIENZE DI I.A. SIMULATA A SCUOLA: RIFLESSIONI DIDATTICHE E NEUROETICHE

I.I.S. Antonello – Messina *– Docenti: Antonio Consolo, Alessandra D'Arrigo, Elisabetta Falduto, Rita Ristagno*

Introduzione

La robotica sociale è quel ramo della robotica che studia l'interazione tra i robot e gli esseri umani.

Una parte della robotica sociale è la Robotica Educativa che studia come l'interazione con i robot possa supportare i processi di apprendimento negli esseri umani. Tradizionalmente gli approcci alla Robotica Educativa sono stati suddivisi in "educazione alla robotica" e "robotica per l'educazione". L'educazione alla robotica, o "apprendimento dei robot", riguarda l'insegnamento tecnico incentrato sulla robotica (gli alunni apprendono il linguaggio di programmazione e come funzionano i robot); mentre la robotica per l'educazione, o "apprendimento con i robot", prevede l'insegnamento di diverse materie (tecniche e non) utilizzando la robotica.

I rapidi progressi nella robotica sociale offrono prospettive promettenti per l'innovazione dell'intervento con i bambini con Bisogni Educativi Speciali. Ad esempio, per quanto riguarda le persone con Disturbo dello Spettro Autistico, alcuni studi suggeriscono che questi soggetti possono interagire più facilmente con un robot piuttosto che con un essere umano. Questo potrebbe essere dovuto principalmente ad alcune caratteristiche dei robot: agiscono come strumenti motivazionali e sono mediatori nella comunicazione, dal momento che il loro comportamento è più prevedibile di quello umano. Con la denominazione di Disturbo dello Spettro Autistico si indicano, infatti, un insieme di disturbi dello sviluppo neurologico caratterizzati da difficoltà nell'interazione sociale e nella comunicazione e da comportamenti e/o interessi limitati e ripetitivi. Le difficoltà maggiori derivano dalla incapacità di leggere le emozioni altrui e, a causa della variegata espressività del volto umano, vi è la difficoltà ad interpretare numerosi segnali sociali come le espressioni facciali, i gesti, il tono della voce.

Si è visto che l'uso di robot umanoidi, cioè di macchine antropomorfe, permette di ridurre al minimo la quantità di informazioni grazie a fattori quali movimenti prevedibili e identici, una voce sintetica con nessuna espressione di personalità e intonazione limitata, la presenza di un software che permette di simulare le abilità sociali e affettive di base. Questa combinazione di caratteristiche in genere facilita la riduzione dell'ansia e migliora la recettività nei giovani con Disturbo dello Spettro Autistico. Tra i diversi robot in commercio, il robot umanoide NAO progettato da Softbank Robotics, rappresenta uno strumento di mediazione terapeutico per questi giovani.

NAO consente inoltre, grazie alla presenza di altoparlanti, sensori tattili, videocamere, microfoni sensoriali, di assumere diverse modalità espressive (movimenti, posture, gesti, voce) e di simulare comportamenti e interazioni intelligenti basati sulla comprensione del linguaggio umano e del linguaggio del corpo, sul riconoscimento delle emozioni e sulla capacità di contatto visivo· Il suo aspetto quasi umano, inoltre, causa un coinvolgimento emotivo da parte di tutti gli studenti, facilitando l'apprendimento, e poiché viene percepito come un gioco, causa un abbassamento dei livelli di ansia e di stress, creando un canale comunicativo spontaneo e automatico.

Il corpo docente della nostra scuola, l'I.I.S. Antonello di Messina, è stato incoraggiato a utilizzare il robot NAO e la Robotica Educativa dalla Dirigente Scolastica Prof.ssa Laura Tringali, che da sempre ha posto particolare attenzione alle esigenze formative di tutti gli

alunni, in un'ottica di strategia inclusiva, tenendo lo sguardo sempre rivolto all'innovazione e in particolare alle sfide e alle opportunità offerte dalle nuove tecnologie.

La convinzione di essere sulla strada giusta ha avuto poi ulteriore conferma dalla letteratura disponibile ad oggi e dalle osservazioni che abbiamo potuto effettuare durante le giornate di Open Day che la nostra scuola ha dedicato all'accoglienza degli alunni delle scuole secondarie di primo grado e delle loro famiglie, per far conoscere la nostra offerta formativa.

La curiosità e lo stupore che la presenza di NAO ha suscitato nei nostri ospiti, insieme alla modalità con cui si sono rapportati e hanno interagito con il robot per cercare di capirne il funzionamento, ha rafforzato la nostra idea sull'importanza di investire tempo ed energie nel suo utilizzo.

Si è così costituito nella nostra scuola un gruppo di lavoro formato da docenti con competenze disciplinari e professionalità diverse, che con entusiasmo si sono messi in gioco per realizzare progetti didattici con la Robotica Educativa. Il lavoro di gruppo ha incentivato la collaborazione e ha consentito di individuare le strategie più efficaci per la risoluzione dei problemi che di volta in volta si sono presentati.

In occasione della Conferenza Nazionale di Bioetica per le scuole 2024 dal titolo "L'etica dell'Intelligenza Artificiale (I.A.): opportunità e rischi" abbiamo deciso, quindi, di partecipare mettendo a frutto tale esperienza e presentando un progetto relativo all'Intelligenza Artificiale Simulata con il robot umanoide NAO come mediatore nella gestione della rabbia.

La nostra esperienza – Fase preliminare

L'idea che ha dato vita al nostro progetto nasce dalla volontà di aiutare un alunno con Bisogni Educativi Speciali a gestire la rabbia.

Da alcune osservazioni effettuate in aula, è emersa la difficoltà dell'alunno a gestire alcune situazioni che gli risultavano particolarmente stressanti; tali situazioni causavano sentimenti di frustrazione con l'insorgenza di collera e rabbia che, pur non comportando gesti violenti, rimanevano incontrollabili. In concomitanza di questi episodi l'alunno manifestava anche il rifiuto a confrontarsi con i docenti della classe, probabilmente per imbarazzo e frustrazione.

Si è progettato, quindi, un intervento educativo che ha previsto, inizialmente la somministrazione di alcune semplici domande in forma di intervista con la compilazione di schede semi-strutturate che hanno consentito di introdurre l'argomento per favorirne la discussione e la riflessione.

Dopo un momento di diffidenza iniziale, piano piano, opportunamente guidato, l'alunno è riuscito ad aprirsi e a confidare le proprie emozioni. Solo a questo punto si è passati all'interazione con NAO.

La nostra esperienza – Fase uno (Educazione alla Robotica)

Abbiamo pensato di realizzare mediante Choregraphe e la programmazione a blocchi, un programma per il robot NAO che potesse suggerire alcune strategie per il contenimento della rabbia,

Per realizzare un intervento efficace e significativo abbiamo cercato di simulare una Intelligenza Artificiale mimando comportamenti e interazioni intelligenti come la comprensione del linguaggio, il contatto visivo, il riconoscimento delle emozioni.

Nella parte iniziale si è cercato di creare un'identificazione dell'alunno con NAO relativamente ai sentimenti di rabbia e frustrazione provata. Per ottenere questo abbiamo dapprima creato un evento scatenante che giustificasse la rabbia "provata" dal robot e, quindi, abbiamo replicato con il robot alcuni atteggiamenti, suoni, gesti messi in atto dall'alunno durante gli episodi di collera. Questo è servito per trasmettere l'idea che alcune

emozioni non riguardano solo noi, ma possono essere vissute anche dagli altri, quindi sono normali.

Nella parte centrale NAO suggerisce una possibile modalità di superamento della rabbia e chiede l'aiuto dell'alunno invitandolo a contare insieme a lui. In questo modo l'alunno si può sentire parte attiva del processo di risoluzione del problema.

Nella parte finale NAO, "ormai calmo", esprime la propria felicità suonando e ballando un breve motivo musicale molto gradito all'alunno e lo ringrazia per l'aiuto ricevuto. L'erogazione di questo tipo di rinforzo positivo di tipo sociale serve a gratificare l'alunno per il successo ottenuto. Il programma termina con NAO che esprime alcune riflessioni riguardo ciò che è accaduto, analizzando le motivazioni che lo hanno portato a perdere la pazienza e offrendo anche un punto di vista diverso per analizzare la situazione.

La nostra esperienza – Fase due (robotica per l'educazione)

Prima di procedere all'interazione tra NAO opportunamente programmato e l'alunno, si sono realizzate alcune sessioni di interazione con il robot, in modalità "consapevolezza di base", in modo da stabilire un primo contatto tra il robot e l'alunno, formulando e rispondendo a semplici domande del tipo "come ti chiami?", "da dove vieni", "quanti anni hai?".

Una volta stabilito un canale comunicativo con il robot, si è passati alla somministrazione del programma realizzato, facendo interagire NAO in modalità autonoma, cioè senza il cavo di rete inserito e senza il controllo diretto di un operatore per mezzo di un PC, così da favorire un maggiore coinvolgimento emotivo da parte dell'alunno.

Alcune manifestazioni spontanee dell'alunno, in risposta agli stimoli forniti dal robot, hanno evidenziato un suo profondo coinvolgimento emotivo: più volte è sembrato riconoscersi e immedesimarsi nelle emozioni simulate del robot, stabilendo con esso un rapporto quasi empatico. L'uso di NAO come mediatore didattico ha permesso di trasferire le emozioni negative su un piano più agevole da condividere dal punto di vista emotivo, in modo da aiutarlo a descrivere i propri sentimenti e a riflettere sulle cause e conseguenze del proprio comportamento.

Questo potrebbe essere il punto di forza da utilizzare in futuro, per favorire quindi il processo di accettazione e contenimento e per fare acquisire una maggiore consapevolezza delle emozioni provate.

Conclusioni

Utilizzare Nao nella pratica educativa consente di creare esperienze didattiche concrete e stimolanti, di sviluppare diverse competenze (come *problem solving*, *cooperative learning*), di implementare la creatività, di realizzare progetti inclusivi. Le attività didattiche laboratoriali favoriscono la motivazione individuale, stimolano le abilità cognitive, emotive e relazionali e promuovono lo sviluppo del pensiero critico.

Il modo in cui NAO si muove, parla, comunica con la gestualità (Intelligenza Artificiale Simulata) lo rende attraente e coinvolgente favorendo l'interazione spontanea, l'attenzione e la concentrazione.

La creazione e l'impiego di *agenti artificiali* come Nao nel settore dell'educazione accendono il dibattito nell'ambito dell'etica e della neuroetica. La tesi della Mente Estesa di Clark e Chalmers, ritiene che «…la mente può realizzare i propri processi estendendosi in *supporti materiali* esterni all'agente» liberandosi «dal confinamento» nel cervello, cioè l'uomo estende la propria mente utilizzando "agenti artificiali" che agiscono per suo conto. Nel nostro caso l'agente artificiale è il robot NAO che, in modalità Intelligenza Artificiale Simulata, è usato come supporto al fine di migliorare l'efficacia della comunicazione e dell'interazione sociale ed emotiva.

Queste attività esperienziali conducono inevitabilmente docenti e studenti a interrogarsi su alcune questioni rilevanti da un punto di vista etico:

- è giusto utilizzare l'Intelligenza Artificiale nei contesti didattico-formativi? E, in caso affermativo, in che maniera?
- e se in futuro l'Intelligenza Artificiale riuscisse ad acquisire una propria coscienza e quindi a comprendere sentimenti e emozioni altrui e a controllarli?
- è giusto utilizzare i robot umanoidi come mediatori sociali con soggetti fragili che subiscono una sorta di inganno nel relazionarsi con il robot come se fosse un altro soggetto umano?
- in un futuro un'I.A. divenuta *embodied* potrebbe riuscire a sfruttare l'interazione empatica con gli esseri umani per condizionarne le emozioni?

La possibilità di generare scambi emozionali simulati tra una Intelligenza Artificiale e gli esseri umani, può presentare dei rischi, in quanto la situazione non è completamente prevedibile e controllabile.

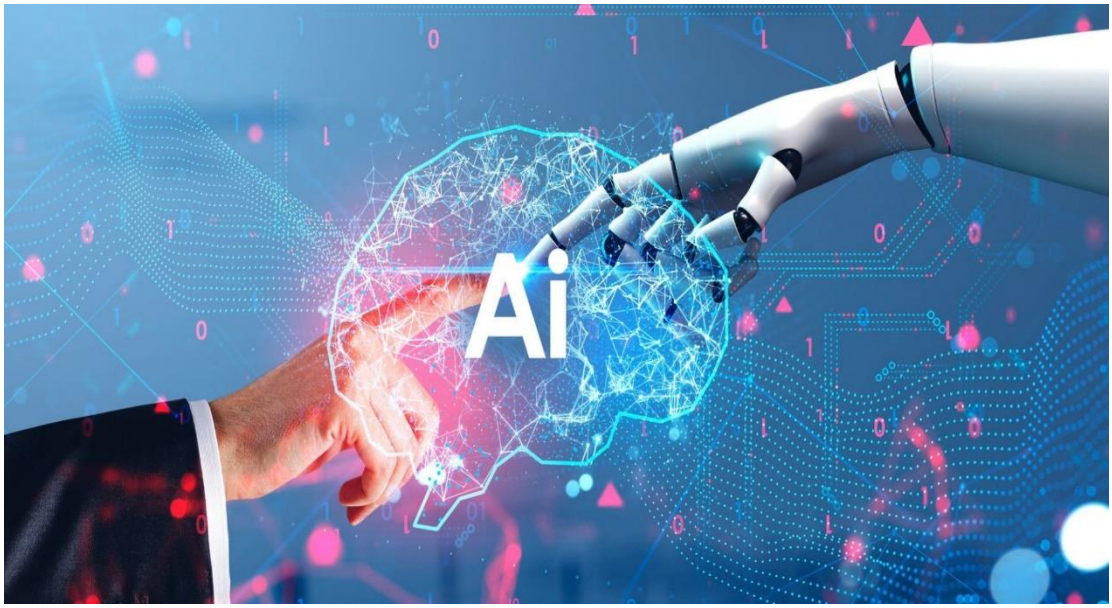
Una possibile soluzione, quella che abbiamo sperimentato direttamente, è di ovviare a tale problema utilizzando una programmazione mirata e sicura, in un ambiente controllato, in modo da realizzare esperienze didattiche in contesti di sicurezza.

Riferimenti Bibliografici

1. M. PIVETTI, S. DI BATTISTA, F. AGATOLIO, et al. "*Educational robotics for children with neurodevelopmental disorders: a systematic review*". Heliyon 6 2020; 10: e05160.
2. M. DUBOIS-SAGE; B. JACQUET; F. JAMET; J. BARATGIN, "*People with Autism Spectrum Disorder Could Interact More Easily with a Robot than with a Human: Reasons and Limits*". Behav. Sci. 2024, 14, 131.
3. COMITATO NAZIONALE PER LA BIOETICA, Disabilità mentale nell'età evolutiva: il caso dell'autismo, 19 aprile 2013, *https://bioetica.governo.it/media/1832/p108_2013_disabilita-mentale-___autismo_it.pdf*
4. American Psychiatric Association, Diagnostic and Statistical Manual of Mental Disorders (fifth ed.), American Psychiatric Pub, Washington, DC (2013)
5. "*Robotica Educativa Autismo: "I robot sociali come aiuto nella terapia per l'autismo*". www.robot.omitech.it, febbraio 2024.
6. "*Nao, a humanoid robot as a therapeutic mediator for young people with autism*". SoftBank Robotics & ERM book, 2021 https://www.aldebaran.com/
7. K. ANTER, M. GREINER, J. VATTER, J. WEGHAKE, DIRECTION: H. SCHNAUBELT. "*Nao6 Learn it, the basic*".
8. G. PALESTRA, B. DE CAROLIS, F. ESPOSITO. *"Artificial Intelligence for Robot-Assisted Treatment of Autism*." *WAIAH@AI*IA* (2017).
9. M. PARENTE, "*La fabbrica dei giochi. Strategie ludiche per bambini con BES*". Erickson, 2010.
10. Choregraphe 2.8.7.4 *https://support.unitedrobotics.group/en/support/solutions/articles/80001018812-nao-6-downloads*
11. A. CLARK, D.J. CHALMERS. "*The extended mind*". In Analysis, 58, 1, pp.10-23, 1998.
12. P. DUMOUCHEL, L. DAMIANO. "*Vivere con i robot. Saggio sull'empatia artificiale*". Raffaello Cortina Editore 2019.

“HEY ALEXA, PRENDERAI IL MIO POSTO?”
INTELLIGENZA ARTIFICIALE E SOSTENIBILITA’

Liceo G. Galilei di Napoli *- Flavia Venosa, Simone Corrado - 4ªB, guidati dalla Prof. Bruna Celano*



Il tema scelto quest’anno dall’Istituto di bioetica è l’utilizzo dell’intelligenza artificiale. La prima parte del nostro lavoro tratta della definizione e della “storia” dell’interazione tra Intelligenza Artificiale e uomo. La seconda parte vuole riallacciarsi al tema, più che mai attuale, della conferenza dello scorso anno sul rapporto uomo-ambiente, e su come l’IA può contribuire alla sostenibilità ambientale.

Prima parte

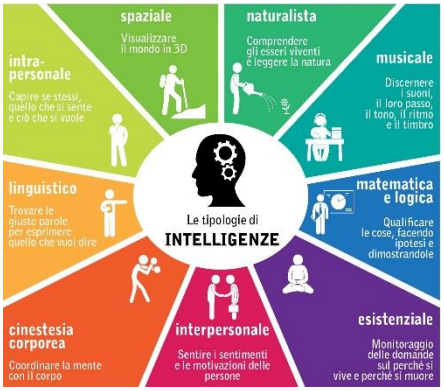


Quando si parla di Intelligenza Artificiale, si pensa subito a tecnologie all'avanguardia, a robot in grado di comprendere e decidere le azioni da compiere e di un mondo futuristico in cui macchine e uomini convivono. In realtà, l'Intelligenza Artificiale e il suo utilizzo sono molto più reali di quanto si possa immaginare e vengono oggi utilizzati in diversi settori della vita quotidiana. L'abbiamo voluto sottolineare nel nostro titolo:

"Alexa, dispositivo che molti possiedono in casa è un assistente vocale dotato di intelligenza artificiale".

Ci pensiamo mai? Probabilmente no ma ciò è dovuto al fatto che non sappiamo cos'è un'intelligenza artificiale.

L'intelligenza artificiale è il settore delle tecnologie dell'informazione e della comunicazione che ha l'obiettivo di imitare con tecnologie informatiche alcuni aspetti dell'intelligenza umana per sviluppare prodotti informatici o macchine in grado di interagire con l'ambiente e l'uomo e assumere decisioni autonomamente. Il termine "Intelligenza" coerentemente con il termine inglese intelligence (tradotto con "informazioni" o servizio di informazioni) non designa qualità propriamente umane conferite alla macchina ma descrive funzioni che rendono alcuni comportamenti delle macchine simili a quelli di un essere umano.



Si tratta cioè, non solo di intelligenza intesa come capacità di calcolo o di conoscenza di dati astratti, ma anche e soprattutto di tutte quelle differenti forme di intelligenza che sono riconosciute dalla teoria di Gardner, e che vanno dall'intelligenza spaziale a quella sociale, da quella cinestetica a quella introspettiva. Un sistema intelligente, infatti, viene realizzato cercando di ricreare una o più di queste differenti forme di intelligenza che, anche se spesso definite come semplicemente umane, in realtà possono essere ricondotte a particolari comportamenti riproducibili da alcune macchine.

Un po' di storia…

A differenza di come potremmo pensare, l'Intelligenza Artificiale nasce con l'avvento dei computer e la sua data di nascita viene fissata nel 1956. Proprio in quest'anno, infatti, si parlò per la prima volta di Intelligenza Artificiale durante un convegno che si tenne in America e che vide la partecipazione di alcuni dei più importanti nomi di quella che sarebbe successivamente stata definita Intelligenza Artificiale, ma che allora veniva denominata Sistema Intelligente. Gli anni successivi furono poi anni di grande fermento intellettuale e sperimentale: università e aziende informatiche, tra cui in particolare l'IBM, puntarono alla ricerca e allo sviluppo di nuovi programmi e software in grado di pensare e agire come gli esseri umani almeno in determinati campi e settori. Nacquero così programmi in grado di dimostrare teoremi sempre più complessi e, soprattutto, nacque il Lisp, ossia il primo linguaggio di programmazione che per oltre trent'anni fu alla base dei software di Intelligenza Artificiale. Da lì in poi si continuarono a cercare software sempre più sofisticati e capaci di risolvere problemi matematici in primis e poi anche ricercare soluzioni a problematiche più vicine alla realtà dell'uomo.

Agli inizi degli anni ’80 il primo sistema di Intelligenza Artificiale fu utilizzato per scopi commerciali e, soprattutto, e gli studi allargarono i propri ambiti geografici, interessando non solo gli Stati Uniti, ma anche il Giappone e l’Europa.

La nuova era dell’I. A. si apre con l’utilizzo di un algoritmo che, già ideato alla fine degli anni Sessanta, non aveva trovato la massima applicazione a causa delle carenze dovute ai sistemi di apprendimento dei primi programmi di Intelligenza Artificiale. Si tratta dell’algoritmo che permetteva l’apprendimento *per reti neurali*, le cui sperimentazioni coprirono sia campi prettamente informatici sia psicologici. Proprio questa doppia applicazione permise agli sviluppatori di Sistemi Intelligenti di trovare un ampio spettro di applicazioni.

In particolare, il primo vero successo dell’Intelligenza Artificiale è stato quello che ha visto il confronto tra *Deep Blue*, una macchina realizzata dalla IBM, e il campione di scacchi allora in carica Garry Kasparov. Anche se i primi incontri furono vinti da Kasparov, i continui miglioramenti apportati al sistema di apprendimento di



Deep Blue permisero, in successive partite, di assicurare la vittoria alla macchina. Una vittoria che, come confermò lo stesso campione di scacchi, fu sicuramente data dal fatto che la macchina
aveva raggiunto un livello di creatività così elevato che andava oltre le conoscenze del giocatore stesso.

É curioso come proprio gli scacchi ci riportano ad oggi: la startup americana Neuralink del miliardario Elon Musk ha pubblicato online un video in cui un paziente, rimasto completamente paralizzato dalle spalle in giù, è riuscito a giocare a scacchi sul proprio computer muovendo il cursore del mouse solamente grazie ad un impianto cerebrale di IA.



Ulteriori applicazioni dei sistemi di Intelligenza Artificiale molto noti al grande pubblico sono quelli utilizzati su veicoli, in grado di guidare senza che vi sia un conducente umano al volante. Si tratta di veicoli ancora in fase sperimentale, ma che raggiungono gradi di sicurezza sempre più elevate soprattutto grazie all'uso di sensori e telecamere che, proprio come occhi e orecchie umane, sono in grado di percepire tutto quanto avviene durante la guida, prendere decisioni ed effettuare manovre di sicurezza.



Uno dei principali passi avanti nella storia dell'Intelligenza Artificiale è stata fatta quando si sono potuti ricreare degli algoritmi specifici, in grado di far migliorare il comportamento della macchina (inteso come capacità di agire e prendere decisioni) che può così imparare tramite l'esperienza, proprio come gli esseri umani.

Sviluppare algoritmi in grado di imparare dai propri errori è fondamentale per realizzare sistemi intelligenti che operano in contesti per i quali i programmatori non possono a priori prevedere tutte le possibilità di sviluppo. Tramite l'apprendimento automatico (*machine learning*), quindi, una macchina è in grado di imparare a svolgere una determinata azione anche se tale azione non è mai stata programmata tra le azioni possibili!

Il *deep learning* è un sottoinsieme del *machine learning*: è essenzialmente una rete neurale con tre o più livelli che tentano di simulare il funzionamento del cervello umano, permettendo al sistema di "imparare" da grandi quantità di dati. Se una rete neurale con un singolo livello è comunque in grado di fare previsioni approssimative, ulteriori livelli "nascosti" possono aiutare a ottimizzare e perfezionare la precisione.

Il deep learning guida molte applicazioni e servizi di intelligenza artificiale che migliorano l'automazione, svolgendo compiti analitici e fisici senza l'intervento umano.

Questo concetto permette di introdurre un argomento molto dibattuto: l'incertezza del futuro, e il timore che l'intelligenza artificiale possa sostituire la capacità decisionale dell'umano, o che la autonomizzazione della macchina possa sfuggire al controllo dell'uomo.

Per quanto teoricamente possibile, siamo ancora lontani da questo scenario: l'I.A. è uno strumento potente ma accessorio alla decisione umana.

Il problema semmai riguarda il fatto che un sistema esperto che diventa ottimale nel suggerire decisioni all'uomo rischia di ridurre l'attenzione umana con la possibile conseguenza della riduzione dell'utilizzo delle abilità umane (*deskilling*). In questo senso è importante una riflessione sulla sinergia tra uomo e macchina e sulla ricerca della modalità di supporto intelligente che consenta all'uomo un controllo significativo in termini di supervisione e attenzione.

Seconda parte

Qual è l'ambito di applicazione dell'IA nelle problematiche ambientali e per la sostenibilità?



L'I.A. e gli IOT (*Internet of things,* cioè l'insieme di oggetti e dispositivi connessi, in grado di ricevere e trasmettere dati da e verso altre reti), possono contribuire notevolmente alla salvaguardia dell'ecosistema, mediante l'acquisizione e l'elaborazione di una vastissima mole di dati provenienti dall'ambiente circostante.

In primo luogo bisogna focalizzarsi su uno dei fattori che maggiormente incrementa fenomeni come quello dell'inquinamento: le industrie. Secondo uno studio condotto da una nota società statunitense, la *Boston consulting group*, l'utilizzo di tecnologie evolute e, in particolare, di sistemi di intelligenza artificiale sarebbe in grado di ridurre fino al 10% le emissioni di gas serra da parte delle aziende entro 10 anni, generando cospicui guadagni, ammontanti a quasi tre milioni di dollari, tra nuovi ricavi e risparmi di costi. Da questo punto di vista possiamo annoverare molteplici funzioni mediante le quali l'intelligenza artificiale può intervenire a favore della sostenibilità ambientale:

-il monitoraggio delle emissioni di gas serra, attraverso la raccolta di dati inerenti alle emissioni della stessa azienda;

-la previsione delle emissioni future, in relazione alle azioni compiute attualmente e che si intendono attuare. (AI predittiva)

-il suggerimento di soluzioni per la riduzione delle emissioni, migliorando l'efficienza energetica delle industrie, dalla produzione alla logistica. (Ai Prescrittiva).

È importante ricordare come le possibilità legate all'impiego dell'IA da un punto di vista ambientale, vadano notevolmente oltre al mero monitoraggio dei consumi aziendali, ma sposino ben più ampie cause.



In primo luogo preservare la biodiversità, ambito in cui un'altra nota società statunitense, la "Blue river technology", sta utilizzando l'IA per rilevare cambiamenti e identificare specie invasive, o "aliene", ossia specie di animali che si ritrovano a vivere in un territorio non corrispondente al loro habitat naturale, la cui diffusione può generare problematiche all'ecosistema e alle attività umane, come nel caso del "granchio blu" nel Mediterraneo, che oltre a arrecare danni alle specie marine ha avuto inevitabilmente effetti negativi sulla pesca determinando la crisi di molte aziende.

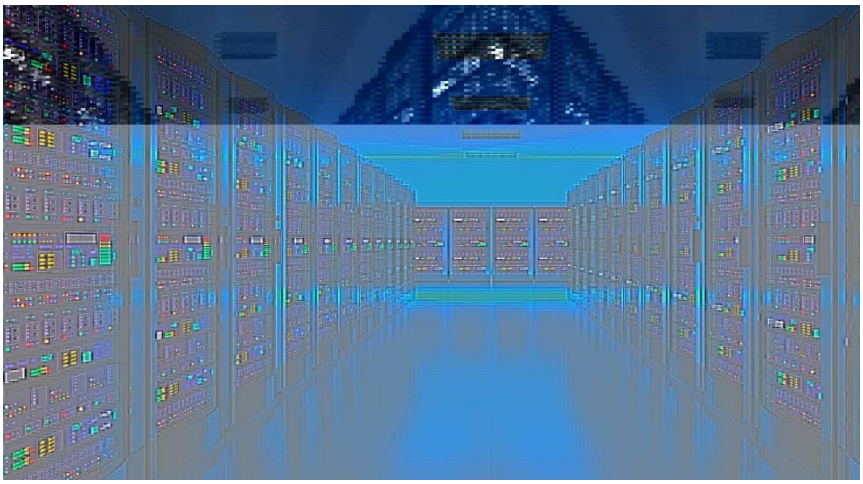
Grazie all'intelligenza artificiale si è in grado di identificare, monitorare e tracciare tali specie, rendendo di fatto più agevole o la loro eliminazione, o la creazione di nuovi equilibri.

Inoltre l'AI è in grado di raccogliere dati provenienti da località oceaniche difficili o impossibili da raggiungere, aiutando così a proteggere specie e habitat. Nello specifico, alcuni robot e IoT, basati sull'AI, possono essere usati per monitorare negli oceani i livelli di inquinamento, di temperatura e di pH, così come condizioni atmosferiche, flora e fauna. Altro collegamento tra AI e IoT può essere quello riguardante l'analisi predittiva dell'IA mediante l'utilizzo di droni, piattaforme di sensori e altri strumenti capaci di monitorare terremoti, inondazioni, tempeste di vento, cambiamenti del livello del mare e altri possibili pericoli naturali affinché si possano attuare azioni tempestive prima che sia troppo tardi.

Esiste, però, inevitabilmente un "rovescio della medaglia," legato ai limiti e ai possibili impatti negativi intrinseci alla natura stessa di tali tecnologie. Difatti se da un lato

l’Intelligenza artificiale è funzionale alla salvaguardia dell’ecosistema, dall’altro sono molteplici i rischi ambientali che si intercorrono proprio con l’uso di tali tecnologie, tanto che la “Royal society” ha evidenziato come il web possa essere considerato il quarto paese più inquinante al mondo!

Nello specifico, l’IA, per quanto fondamentale nel monitoraggio e nella riduzione dell’inquinamento industriale, è allo stesso tempo paradossalmente responsabile di una significativa quantità di emissioni di gas serra, e di un notevole consumo a livello energetico e idrico direttamente proporzionale alla potenza di calcolo. Difatti i sistemi di IA per operare necessitano di *data center,* ossia enormi stanze che ospitano



server e apparecchiature necessarie al loro funzionamento, per il cui raffreddamento sono necessarie ingenti quantità di acqua (basti pensare ai 16 mlrd di litri usati da google nel 2021) e la cui richiesta energetica tenderà a duplicarsi nel corso di 7 anni.

A tale proposito uno studio del dipartimento di informatica dell’università del Maryland ha calcolato l’impatto dell’utilizzo di *chat-gpt* nell’arco di una singola giornata, approssimando per difetto, dimostrando come esso corrisponda, in termini di consumo energetico e di emissioni di CO_2, a quello che in media un individuo genera in più di tre mesi, permettendoci quindi di affermare generalmente, che a una maggiore complessità algoritmica corrisponda un maggiore impatto a livello ambientale!

A questo punto bisognerebbe tirare le somme per constatare se il bilancio effettivo è positivo o negativo.

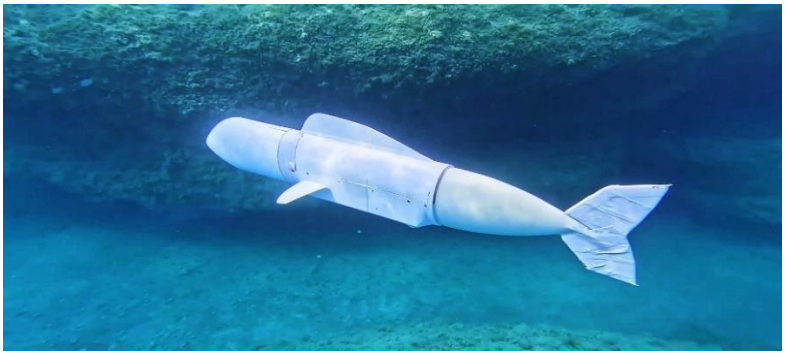
Nonostante l’IA possa incrementare fenomeni come l’inquinamento e di conseguenza minare la sostenibilità, non dobbiamo assolutamente abbandonare la sua applicazione, bensì ricercare un compromesso: sviluppare soluzioni sostenibili che abbiano basso impatto ambientale. Risulta quindi necessario investire in tecnologie di efficienza energetica per i data center e per i sistemi di raffreddamento compatibili con tecnologie sempre più avanzate dell'Al, in modo da poter incentivare lo sviluppo in modo sostenibile. In tale ottica, tra le più significative esemplificazioni, possiamo annoverare gli interventi delle note aziende Google e Meta, che hanno recentemente realizzato *datacenter* in luoghi molto freddi per garantire minori consumi di energie.

Nello specifico Google ha ristrutturato una vecchia cartiera in Finlandia, sfruttando la vicinanza del golfo per avere accesso agevolmente a acqua a bassa temperatura. Meta invece ha costruito un data center in Svezia, dove si utilizzano esclusivamente energie rinnovabili, e dove il clima artico permette, anche in questo caso, una riduzione dell’energia necessaria per il raffreddamento dei server.


Google



Concludiamo con uno dei più singolari esempi di collaborazione tra *IoT* e A.I., finalizzato alla salvaguardia dell'ambiente e in particolare della biodiversità marina: BELLE, il pesce robot sottomarino, autonomo, dotato di telecamera e progettato per raccogliere campioni di DNA sottomarini per la ricerca marina. Questo affascinante progetto, elaborato dagli studenti di ingegneria meccanica di Zurigo, è frutto di



un'innovativa idea, quella di realizzare una struttura che potesse amalgamarsi e confondersi perfettamente nell'ecosistema, venendo accettata come parte integrante di esso. Ciò ha condotto alla progettazione di un robot che si comporta come un pesce e viene accettato dalle altre creature marine come tale. Il robot è lungo 3 piedi, utilizza l'intelligenza artificiale per muoversi autonomamente mediante l'ausilio di una pinna in silicone con due cavità che pompano acqua in cicli mentre la sua coda morbida si solleva e si abbassa, non creando nessuna turbolenza nell'acqua e permettendo al robot di mimetizzarsi perfettamente con gli altri pesci negli ambienti sottomarini



Belle, mediante la sua piccola telecamera può filmare sott'acqua video ad alta risoluzione e raccogliere campioni che forniscono una mappatura dettagliata degli organismi che vivono in un determinato ambiente, rendendo agevole l'eventuale individuazione di specie invasive o "aliene" e di conseguenza agevolarne l'eliminazione o il monitoraggio, fornendo dunque significativi vantaggi per preservare la biodiversità.

INTELLIGENZA ARTIFICIALE E PRIVACY: UN INCONTRO POSSIBILE?

Liceo P. Gobetti Genova. *Ben Hassine Jonathan, Cosso Samuele, Foli Diletta - Classe V B SU,*

Il nostro intervento si propone di indagare sulle possibilità offerte dall'utilizzo dell'intelligenza artificiale ma anche sui nodi problematici portati da un uso non consapevole dell'intelligenza artificiale.

Per effettuare un'analisi il più possibile esaustiva abbiamo cercato di fornire una definizione di intelligenza artificiale, di scoprire le connessioni tra intelligenza artificiale e neurotecnologie per arrivare a comprendere come sia possibile l'utilizzo consapevole di queste tecnologie, capaci di modificare profondamente la nostra vita.



Nella letteratura scientifica non esiste una definizione di IA univoca. Questo perché, nel corso degli anni, sotto l'espressione intelligenza artificiale sono stati raccolti contributi provenienti da campi scientifici e tecnologici diversi, rendendola una disciplina di confine tra scienza e ingegneria.

Inoltre, la presenza della parola intelligenza, già di per sé complessa da definire, complica ulteriormente la questione. "*L'IA è quella disciplina che studia e costruisce algoritmi che cercano di rendere i computer in grado di compiere azioni che richiederebbero facoltà tipiche della mente*"[7].

I sistemi di IA, quindi, sono capaci di adattare il proprio comportamento, analizzando gli effetti delle azioni precedenti e lavorando in "autonomia". Le neurotecnologie sono dispositivi in grado di registrare, decodificare e/o modificare l'attività del cervello.

L'incontro tra IA e neurotecnologie può produrre moltissimi vantaggi, ad esempio nella medicina di precisione, nella scoperta di nuovi farmaci, nella cura e prevenzione delle malattie, nel rendere maggiormente efficiente il sistema di salute pubblica.

Le neurotecnologie sono tecnologie che hanno il potere di trasformare radicalmente le nostre vite, quindi rappresentano una vera e propria sfida che l'umanità dovrà affrontare nel prossimo futuro. Per questo è giunta l'ora di porsi il problema della tutela della nostra privacy, quella che coinvolge i nostri pensieri e le nostre emozioni più profonde.

Coloro che intendono applicare la IA insistono sulla necessità di disporre di dati in un vasto ambito di azione, sul piano globale (dunque anche trasferimento di dati in altre nazioni)

[7] G. Savino, *Intelligenza artificiale, tra timori e reali possibilità.* In Aula di Scienze- persone storie e dati per capire il mondo, 11 gennaio 2024, pag. 5.

e sulla conservazione dei dati, ma anche di campioni e immagini associate nel tempo. Dati non pienamente anonimi ma pseudonimizzati che consentano una tracciabilità[8].

Da quanto scritto risulta fondamentale un utilizzo responsabile dell'intelligenza artificiale al fine di tutelare la nostra privacy, ma come raggiungere questo obiettivo? Come garantire ai cittadini la sovranità sulla propria mente?

Innanzitutto, è importante permettere a ciascun individuo l'accesso alle informazioni sulla propria salute e sul proprio benessere in modo che possa assumerne il controllo. Le organizzazioni inoltre dovrebbero implementare misure solide per salvaguardare le informazioni sensibili, aderire alle normative sulla protezione dei dati e dare priorità al consenso e al controllo degli utenti sui propri dati.



Il trattamento responsabile dei dati garantisce la fiducia degli utenti e mantiene gli standard etici. Nella nostra società digitale, in rapida evoluzione, la privacy rappresenta il diritto fondamentale alla riservatezza e al controllo dei dati personali, una risorsa preziosa in un mondo sempre più interconnesso.

La crescente raccolta di analisi e utilizzo dei dati da parte degli algoritmi di IA pone sfide significative alla protezione della privacy, richiedendo una riflessione attenta e misure efficaci per garantire che i diritti individuali siano rispettati. La raccolta dei dati è il primo passo: con l'avvento di tecnologie sempre più avanzate, una vasta gamma di informazioni personali viene costantemente acquisita da diverse fonti, ad esempio dai dispositivi, dai social media, dalle transazioni online e altro ancora. Questi dati, una volta raccolti, vengono spesso aggregati e analizzati da algoritmi di IA per estrarre informazioni significative e modelli comportamentali.

L'analisi dei dati personali tramite algoritmi di IA può rivelare dettagli intimi sulla vita delle persone, inclusi gusti personali, abitudini di consumo, relazioni sociali e persino informazioni mediche sensibili. Questo processo di analisi può essere estremamente invasivo e sollevare preoccupazioni sulla privacy, specialmente quando i dati vengono utilizzati senza il consenso esplicito dell'utente o per scopi non autorizzati.

Gli individui si trovano, quindi, in un evidente stato di vulnerabilità; perciò, è essenziale salvaguardare efficacemente la loro libertà di pensiero: una libertà che andrebbe ampliata, includendo la libertà cognitiva, intesa come il diritto all'autodeterminazione per quanto riguarda il cervello e le esperienze mentali.

[8] Cfr., L. Battaglia, *I potenziali e i pericoli delle neurotecnologie*, il secolo XIX, 10 marzo, 2024.
Cfr. cfr., G. Collecchia, *Neurotecnologie: verso la lettura del pensiero?,* 2 maggio 2021, in https://www.meer.com/it/65654-neurotecnologie-verso-la-lettura-del-pensiero



È indispensabile, quindi, fornire gli strumenti necessari per definire un nuovo diritto alla libertà cognitiva, che comprende i diritti alla privacy mentale alla libertà di pensiero, all'autodeterminazione, perché l'utilizzo dei dati personali, da parte dei sistemi di IA, può avere conseguenze significative sulla privacy e conseguentemente sul nostro ben-essere e ben-vivere: le decisioni e le raccomandazioni generate dagli algoritmi di IA possono aumentare le opportunità di lavoro, l'accesso ai servizi finanziari, le possibilità di assicurazione e molto altro ancora. Tuttavia, se questi algoritmi non sono adeguatamente controllati e regolamentati, c'è il rischio che possano perpetuare discriminazioni o pregiudizi, compromettendo ulteriormente la privacy e la giustizia sociale.

In questa prospettiva è fondamentale il ruolo dell'educazione: innanzitutto, è importante educare gli individui, comunicando i vantaggi e i rischi portati dall'utilizzo dell'IA, spiegando i concetti di base, le applicazioni e le potenziali implicazioni etiche la libertà e la privacy. Questo implica un'adeguata informazione sul funzionamento degli algoritmi, sulla raccolta e sull'analisi dei dati, nonché sulle possibili conseguenze delle "decisioni" prese dagli algoritmi.

La sensibilizzazione pubblica svolge un ruolo complementare, diffondendo la consapevolezza sulle questioni di privacy ed etica nell' utilizzo dell'IA, attraverso campagne informative, eventi educativi e materiali divulgativi. Questo può contribuire a creare un dibattito pubblico informato e stimolare l'interesse verso utilizzi responsabili e sostenibili dell'IA[9].

Una maggiore consapevolezza e partecipazione pubblica possono influenzare le decisioni dei legislatori, incoraggiando la creazione di leggi e regolamenti che proteggano la privacy e promuovano una riflessione etica capace di valutare i costi e i benefici dell'utilizzo dell'IA. Infatti, nel mese di marzo 2024 Il Parlamento europeo approva la legge sull'intelligenza artificiale al fine di dare:

1. garanzie per i sistemi di intelligenza artificiale usati per finalità generali;
2. limiti all'uso dei sistemi di identificazione biometrica da parte delle forze dell'ordine;
3. rifiutare i sistemi di credito sociale, o per manipolare e sfruttare le vulnerabilità degli utenti[10].

In conclusione: l'IA ha il potenziale per migliorare molti aspetti della nostra vita, ma è essenziale adottare misure rigorose per proteggere la privacy durante la raccolta, l'analisi e l'utilizzo dei dati personali, garantendo che i diritti individuali siano sempre rispettati e preservati.

[9] Cfr., *https://bioetica.governo.it/media/vkkphteb/intelligenza-artificiale-e-medicina.pdf*, Presidenza del Consiglio dei ministri, 29 maggio 2020.

[10] Cfr. Il regolamento UE in materia di intelligenza artificiale n. 26 5 febbraio 2024.

Bibliografia e sitografia

1. G. Savino, *Intelligenza artificiale, tra timori e reali possibilità.* In Aula di Scienze- persone storie e dati per capire il mondo, 11 gennaio 2024.
2. L. Battaglia, *i potenziali e i pericoli delle neurotecnologie*, il secolo XIX, 10 marzo, 2024.
3. G. Collecchia, *Neurotecnologie: verso la lettura del pensiero?,* 2 maggio 2021, in https://www.meer.com/it/65654-neurotecnologie-verso-la-lettura-del-pensiero.
4. *https://bioetica.governo.it/media/vkkphteb/intelligenza-artificiale-e-medicina.pdf*, Presidenza del Consiglio dei ministri, 29 maggio 2020.
5. Il regolamento UE in materia di intelligenza artificiale n. 26 5 febbraio 2024.

L'ERA DELL'INTELLIGENZA ARTIFICIALE TRA SCIENZA E SOCIETÀ

Liceo Pertini*, *GENOVA *II I SC - Docente: Rosaria Catanoso*
Relatori: Angelica Baldassini, Lorenzo Bernardinelli, Carola Carruba, Ginevra De Paoli, Luca Cuomo, Eleonora Sobrero

L'intelligenza artificiale: rischi e prospettive

L'origine della ricerca

L'intelligenza artificiale ha dominato i titoli dei giornali recenti con le sue promesse e con le sue sfide, con i suoi rischi, con i successi e con i fallimenti. Quale futuro possiamo prevedere per l'I.A.? Naturalmente, le previsioni più accurate vengono effettuate con il senno di poi. Eppure, noi oggi, viviamo già immersi nell'era dell'intelligenza artificiale[11].

Il seguente contributo vuol mostrare come le scienze umane[12], le scienze naturali, il diritto[13], l'informatica[14], la cybernetica[15], le neuroscienze siano ambiti di indagine legati proprio dal tema afferente all'intelligenza artificiale[16].

Abbiamo discusso, abbiamo analizzato, ci siamo informati, abbiamo fatto ricerca. Successivamente sono sorte molte domande critiche sul nostro coinvolgimento.

Siamo incuriositi ed interessati perché percepiamo l'intelligenza artificiale presente nella nostra vita e nella nostra quotidianità. Ecco alcune scottanti questioni emerse, durante la fase del dibattito. Che riscontro ha l'intelligenza artificiale nella nostra vita? Qual è il suo obiettivo? Perché e come ci coinvolge? Perché e come ci influenza? Quali ambiti della vita sono coinvolti? Quali sono i suoi vantaggi? Quali gli svantaggi? Perché questi ultimi ci inquietano?

Il punto d'inizio della nostra analisi è stato, proprio, di natura storica e scientifica, seguendo il percorso biografico e l'evoluzione degli studi posti da Alan Turing, per cogliere come già durante la Seconda Guerra Mondiale si volesse pensare un sistema utile a sconfiggere i nazisti, su un terreno tecnologico ed artificiale[17].

Perché ci interessa?

Il primo punto dal quale analizzare il problema posto dall'uso dell'intelligenza artificiale è tentare di capire le linee di indagini successive, i risvolti ed i percorsi di ricerca successivi, considerando il suo stato attuale. Su questa base, lo scopo è provare a tracciare la mappa delle sfide etiche da prendere sul serio.

Esordiamo con un'analisi semantica ed ermeneutica: nella letteratura scientifica non esiste una definizione di IA univoca. Nel corso degli anni, sotto l'espressione intelligenza artificiale sono stati raccolti contributi provenienti da campi scientifici e tecnologici diversi, rendendola una disciplina di confine tra scienza e ingegneria. Inoltre, la presenza della parola intelligenza complica ulteriormente la questione. L'intelligenza artificiale cerca di

[11] R. BODEI, *Dominio e sottomissione. Schiavi, animali, macchine, Intelligenza Artificiale*, il Mulino, Bologna 2019.
[12] E. DATTERI, *Che cos'è la scienza cognitiva*, Carocci, Roma 2017.
[13] M. C. NUSSBAUM, *Le nuove frontiere della giustizia*, il Mulino, Bologna 2007.
[14] M. DAVIS, *Il calcolatore universale. Da Leibniz a Turing,* Milano, Adelphi 2003.
[15] G. TAMBURRINI, *I matematici e le macchine intelligenti. Spiegazione e unificazione nella scienza cognitiva*, Mondadori, Milano 2002.
[16] S. SCAGLIARINI (a cura di), *Smart roads e driverless cars: tra diritto, tecnologie, etica pubblica,* Giappichelli, Torino 2019.
[17] H. HOSNI, (a cura di) *Menti e macchine. Alan Mathison Turing a cento anni dalla nascita,* Edizioni della Normale, Pisa 2015.

mettere in grado i computer di fare il genere di cose che sanno fare le menti; queste azioni coinvolgono capacità psicologiche come la percezione, l'associazione, la previsione e il controllo motorio, consentendo agli uomini di conseguire i propri obiettivi[18].

Iniziamo, quindi, con l'osservare come la normativa europea voglia garantire che le applicazioni di IA siano in linea con aspettative socialmente accettabili.

Ecco perché l'Unione Europea ha sancito la *Proposta di regolamento del Parlamento e del Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale* (COM (2021)206 definitivo[19]. Lo scopo della Proposta è uniformare la ricerca, proprio perché non tutti i Paesi europei sembrano essere, allo stato attuale, universalmente accomunati dal medesimo fine. Del resto, fuori dall'UE, lo sviluppo delle nuove tecnologie non è regolamentato. Ricordiamo come l'Arabia, la Cina, il Giappone, l'India sono alcuni dei Paesi nei quali le ricerche non mostrano limiti normativi; al punto tale che in questi territori le ricerche spaziano su campi di indagine nuovi ed inarrestabili. Temiamo, non tanto che altri Paesi siano un passo avanti nella ricerca scientifica e tecnologica, quanto non sapere cosa possano inventare; ci inquietano gli scopi e gli usi di questi progressi.

Aiuto o sconfitta del nostro fare?

L'obiettivo dell'intelligenza artificiale è di aiutarci a svolgere le azioni che compiamo quotidianamente, rendendole più semplici e sicure; per eseguire questi compiti con successo, non deve per forza essere intelligente. Del resto, esecuzione ed intelligenza non si equivalgono direttamente. Le applicazioni del mio cellulare non devono essere intelligenti, ma solo operative. Molti dei problemi che cerchiamo di risolvere con l'intelligenza artificiale si verificano nel mondo fisico. Pensiamo alla guida, alla scansione di etichette in un supermercato, alla pulizia dei pavimenti, al taglio dell'erba in giardino. In particolare, le abilità di scrittura di un prodotto software hanno sconvolto anche gli ambiti di apprendimento e la scuola; il prodotto in questione è ChatGPT, un assistente virtuale interattivo al quale si può chiedere di tutto tramite una chat.

ChatGPT dialoga fluentemente nelle lingue più diffuse e ha proprietà di linguaggio e comprensione del contesto che rendono le conversazioni fruttuose e piacevoli. In realtà, chiunque usi uno smartphone ha già avuto a che fare, più o meno consapevolmente, con assistenti virtuali basati su tecniche di IA, basti pensare a Siri di Apple o a Google Assistant. ChatGPT è migliore di tutti gli altri assistenti prodotti finora; le conversazioni con ChatGPT sono il risultato di un lavoro di ricerca e sviluppo tecnologico al confine tra diverse branche dello scibile umano come la matematica, l'informatica e l'ingegneria.

Comprendere gli sviluppi dell'intelligenza artificiale, in relazione ad ambienti analogici e digitali, è utile per analizzare le risorse necessarie. Pensiamo alle risorse computazionali, ai gradi di complessità, e alle risorse relative alle abilità e ai gradi difficoltà. Ogni volta che la scissione tra esecuzione ed intelligenza è realizzabile diventa possibile qualunque soluzione di intelligenza artificiale. Questo è il motivo per cui comprendere il futuro dell'intelligenza artificiale significa, anche, cogliere la natura dei problemi in relazione ai quali tale scissione può essere tecnicamente realizzabile, almeno in teoria, ed economicamente attuabile in pratica.

Nel 2018, DeepMind, in collaborazione con l'ospedale Moorfields Eye di Londra, ha addestrato un sistema di I.A. per identificare sintomi di malattie degli occhi pericolose per la vista utilizzando i dati della tomografia a coerenza ottica, una tecnica di elaborazione di immagini che genera immagini 3D della parte posteriore dell'occhio. Questo episodio mostra come l'IA abbia maggiori possibilità di successo ogni volta in cui i sistemi di dati ben curati, aggiornati e completamente affidabili saranno disponibili ed accessibili per addestrare un

[18] M.A. Boden, *L'intelligenza artificiale*, il Mulino, Bologna 2019.
[19] L. Floridi, *Etica dell'intelligenza artificiale. Sviluppi, opportunità, sfide, Raffaello Cortina*, Milano, 2022, p.66.

sistema in un'area specifica di applicazione. L'I.A. diventa uno strumento per migliorare la vita delle persone.

Umanoidi ed ultime frontiere

Le ultime frontiere dell'intelligenza artificiale sconvolgono e scuotono, proprio perché ci troviamo a doverci confrontare umanoidi sensibili, quasi quanto gli umani. Lo scopo delle ultime ricerche, cui stiamo assistendo, è quello di giungere ad impiantare il cervello umano in un robot. Altra intenzione è quella di provare a creare artificialmente le emozioni. L'I.A. può provare emozioni?

Diventa eclatante come in Arabia Sara, un robot, accolga all'ingresso di un museo i visitatori; in Giappone nella reception di un hotel, i clienti trovano un umanoide ad accoglierli. Suscita sgomento, quasi angoscia, vedere la giornalista del tg1 al cospetto del suo umanoide Erica.

La somiglianza con la figura umana sta giungendo ad un realismo rappresentativo impensato, producendo un brusco calo delle reazioni emotive positive[20]. Alla vista di un umanoide, infatti, abbiamo provato disgusto, repulsione, inquietudine, turbamento. Ecco perché siamo a disagio innanzi alla possibilità di un altro noi difficilmente definibile, ma con una sensibilità quasi pari alla nostra.

I.A. nella vita quotidiana

Nel 2016 la Nike ha lanciato Hyper-Adabpt 1.0, le sue scarpe automatiche elettroniche che si allacciano da sole, reinventando un concetto di cosa significa adattare le scarpe ai piedi. Ogni scarpa ha un sensore, una batteria, un motore e un sistema di cavi che insieme possono regolare l'adattamento tramite un algoritmo di pressione. Succedono cose strane quando il software non funziona correttamente.

Secondo la nostra opinione vi sono diversi lati positivi e negativi riguardo questa creazione. Pensiamo alle complicazioni che possono sorgere sul software della scarpa, quelle sono – di certo – negative. In realtà, pensata in ambito sportivo, durante una partita di basket, una tale scarpa renderebbe più sicuro il movimento e il gioco fino al termine della partita. Gli sportivi, infatti, si sentono sicuri con una scarpa da non allacciare e che tutela i movimenti del piede; Da un punto di vista educativo, nessun insegnante o genitore deve consentire che un bambino indossi una scarpa con tali sensori; la motricità fine, l'uso delle dita e l'acquisizione di gesti e movimenti è importantissimo per lo sviluppo e la crescita.

Eseguire un compito è lecito, sostituire questo oggetto pensante è controproducente. A riguardo, soffermiamoci su quanto accaduto in tempi che ormai consideriamo lontani; nel 2012 un team dell'università Carlos terzo di Madrid, in Spagna, ha costruito TEO, un robot di circa 80 kili alto 1.80.

Teo può salire le scale, aprire le porte e stirare. Questo è il futuro? Quali sono i compiti difficili e complessi e l'enorme vantaggio di svolgere i compiti per renderli facili? Non dobbiamo cercare di imitare gli esseri umani attraverso l'IA. Dobbiamo sfruttare ciò che le macchine, inclusa l'IA, fanno meglio.

Il forno proposto da una nota casa di elettrodomestici consente per mezzo di una telecamera interna di controllare la cottura a distanza dal cellulare; può inoltre essere collegato ad Alexa. Queste funzioni, forse, semplificano la vita, non dovendo più controllare la cottura con lo stuzzicadenti, ed evitano che i cibi vengano dimenticati o bruciati, nel caso in cui stessimo svolgendo altre attività.

Cosa pensiamo? Riflettiamo insieme. Una nota casa farmaceutica ha inventato Pillo, un robot che dà al paziente la medicina giusta al momento opportuno.

[20] G. Tamburrini, *Etica delle macchine. Dilemmi morali per robotica e intelligenza artificiale*, Carocci, Milano 2020.

Poniamoci, ancora, una domanda queste invenzioni ci aiutano davvero?

In alcuni casi, sì. Possono fornire un supporto alle persone con disabilità. Invece, in altri casi, queste invenzioni servono a semplificarci la vita, e non possiamo ritenerle utili da ogni punto di vista[21].

Dobbiamo evitare che l'intelligenza artificiale faccia tutto al posto nostro. Dobbiamo sviluppare le nostre capacità e adoperarle. Il cinema, in merito, offre un esempio divertente proprio per mezzo della storia proposta dalla PIXAR dal titolo Wall-E.

Come ci coinvolge?

L'IA ci coinvolge perché quotidianamente adoperiamo un cellulare, chiediamo a Siri di fare cose al posto nostro, parliamo con Alexa, controlliamo la nostra casa da remoto, già con queste piccole cose interagiamo con forme accettabili di I.A.

L'I.A. influenza le nostre vite in vari modi, migliorando i servizi online, cambiando il modo in cui interagiamo con la tecnologia. Inoltre, può influenzare le nostre decisioni quotidiane attraverso raccomandazioni personalizzate e può generare preoccupazioni sull'automazione di posti di lavoro. Quindi, è importante comprendere come funziona per poter gestire in modo efficace le sue implicazioni sulla società e sfruttarne i benefici.

Gli ambiti della vita coinvolti, in altre parole, sono molti quello lavorativo, quello sanitario e medico, con la chirurgia a distanza, l'ambito bellico[22] che desta ansie e grossi timori.

Il futuro dell'I.A.

Il futuro dell'I.A. è pieno di incognite sconosciute. Il futuro prevedibile dell'IA dipenderà dalle nostre capacità di ingegno. Dipenderà anche dalle nostre capacità di negoziare le derivanti e serie questioni etiche, giuridiche e sociali dalle nuove forme di privacy, alle forme di autodeterminazione. L'idea stessa di plasmare sempre più i nostri ambienti analogici e digitali per renderli adatti all'IA deve far riflettere. Prevedere tali questioni, per facilitarne la soluzione ed evitarne o mitigarne gli aspetti negativi, è il vero valore di qualsiasi analisi rivolta al futuro. Il futuro non è del tutto aperto, perché il passato lo plasma, ma non è neppure del tutto determinato, perché il passato può essere indirizzato in una direzione diversa. Ecco perché la sfida che ci attende non sarà quella dell'innovazione, quanto piuttosto della governance del digitale, compresa dell'IA.

L'IA non è l'ultima panacea, che avrebbe risolto e superato tutto. Non è una super-intelligenza che avrebbe distrutto milioni di posti di lavoro, sostituendo avvocati e medici, giornalisti e ricercatori, camionisti e tassisti, e finendo per dominare gli esseri umani come se fossero animali domestici. Tutto dipende, sempre, da noi. Dalle nostre scelte e dalle nostre decisioni. Ecco quindi, come sconfiggere, anche, le possibili paure legate all'impossibilità di controllare gli esiti dei processi innescati dall'I.A; come ben mostrato dal noto film della Marvel in cui i protagonisti si scontrano con ULTRON, una forma di I.A. creata in modo scorretto, inconsapevole ed improprio. L'I.A. non deve in ultimo, quindi, mai e poi mai sostituirci. Noi rimaniamo i protagonisti della nostra vita.

[21] V.S., *Automobili a guida autonoma. Un'inedita opportunità per le persone con disabilità fisiche*? in S. Scagliarini (a cura di), *Smart roads e driverless cars. Tra diritto, tecnologie, etica pubblica*, cit., pp. 55-64.

[22] G. Bangone, *La guerra al tempo dei droni. Da Falluja ai terroristi dell'Isis: la nuova frontiera dei conflitti,* Castelvecchi, Roma 2014.

LA POTENZA DI CONTROLLO DELL'IA SULLA VITA E SULLE IDEE DEI CITTADINI

ITI Marconi – Pontedera *- A cura del Centro di Bioetica di Pontedera (CBP)*
Studenti: Campana Valerio (5AE), Comanciu Razvan (4BI), Marianelli Manuele (3BI), Mazzanti Giada (4BSA), Meloni Alessandro (4A SA), Pratali Matteo (3BTE), Scarpellini Giulio (4BI), Schimizzi Mattia (4ASA), Senesi Marco (4BI), Testi Guido (3BI), Uka Virginia (3AGC) Docenti: Dimitri Bigoni, Antonio Di Giorgio e Simone Matteoli.

ABSTRACT

La potenza di controllo dell'IA sulla vita e sulle idee dei cittadini è un tema cruciale affrontato dal Centro di Bioetica di Pontedera – dell'ITIS G. Marconi - per la Conferenza Nazionale di bioetica per le Scuole svoltasi il 9 aprile 2024.

In questo documento, sono state analizzate varie aree di impatto dell'intelligenza artificiale (IA) nella società, con un focus su tre principali ambiti: l'uso dell'IA nei social media e negli assistenti digitali, i sistemi di credito sociale e l'impatto dell'IA sui neurodiritti.

La prima parte del documento esamina come l'IA sia impiegata nei social network per gestire la pubblicità, analizzare contenuti e stimolare dipendenze digitali attraverso la personalizzazione e il targeting. Si discute anche del ruolo degli assistenti digitali come raccoglitori di dati personali, sollevando interrogativi sull'ottenimento del consenso degli utenti.

Nella seconda parte, vengono esplorati i sistemi di credito sociale, con particolare attenzione al caso cinese e alle implicazioni dell'IA nel controllo sociale attraverso la valutazione dei comportamenti individuali.

Infine, la terza parte affronta il tema dei neurodiritti, analizzando le tecnologie emergenti che consentono la lettura del pensiero e il ruolo chiave dell'IA nelle nanotecnologie. Si discute la necessità di nuovi diritti per proteggere la privacy e l'autonomia individuale di fronte alle avanzate capacità dell'IA nel manipolare e interpretare i processi mentali. Inoltre, si esamina il sistema di credito sociale cinese, che solleva domande significative riguardo alla privacy e alla libertà individuale. Sebbene questo sistema possa fornire incentivi per comportamenti socialmente desiderabili, il rischio di controllo eccessivo da parte dello stato e di violazioni della privacy è notevole. L'analisi tocca anche l'uso di sistemi di sorveglianza di massa e l'integrazione di tecnologie avanzate come il riconoscimento facciale, che sollevano preoccupazioni per i diritti individuali e la libertà civile.

Per quanto riguarda le tecnologie neurali, l'IA ha rivoluzionato il modo in cui possiamo interagire con il cervello umano. Se da un lato queste tecnologie offrono possibilità di miglioramento della qualità della vita per persone con disabilità o condizioni mediche, dall'altro sollevano importanti questioni etiche riguardanti la privacy, la manipolazione mentale e l'autonomia individuale.

L'analisi dei neurodiritti sottolinea l'importanza di proteggere gli individui dall'uso improprio delle neurotecnologie e dell'IA. La definizione di diritti come la privacy mentale, la continuità psicologica, l'integrità mentale e la libertà cognitiva è cruciale per garantire che le persone mantengano il controllo sui propri pensieri e processi mentali, e che vengano rispettati i loro diritti fondamentali.

PARTE 1: IA, social network e assistenti digitali

Utilizzo dell'IA da parte dei Social Network

Gli strumenti basati sull'IA aiutano a potenziare le funzionalità dei social media e ad effettuare operazioni su larga scala in vari contesti, come:

- la produzione di contenuti scritti e visivi
- il monitoraggio dei social media
- la gestione della pubblicità
- la targhettizzazione dei contenuti.

Fra questi uno dei campi più importanti dei social media, per il quale viene utilizzata l'IA, riguarda la gestione delle pubblicità, dove viene usata la tecnica della Computer Vision.

IA e Computer Vision: La lettura delle nostre immagini e video

L'IA può analizzare i comportamenti e le preferenze degli utenti per creare una *sequenza personalizzata di post*. Registra, ad esempio, i video guardati o ai quali si è messo like e propone qualcosa di simile. Analizza i contenuti pubblicati. Questo livello di personalizzazione è particolarmente utile nel marketing sui social media, dove l'obiettivo è coinvolgere e interagire con il pubblico su un livello personale.

Alcuni dettagli tecnici. La visione artificiale, o *computer vision*, è un campo dell'intelligenza artificiale che permette ai computer di ricavare informazioni significative da immagini digitali, video e altri input visivi e intraprendere azioni o formulare delle segnalazioni sulla base di tali informazioni.

La *computer vision* funziona più o meno come la vista umana, ma la vista umana ha il vantaggio di disporre di anni di esperienza in cui si è allenata a distinguere gli oggetti, quanto sono lontani, se si stanno muovendo e se c'è qualcosa di sbagliato in un'immagine.

Questa ricerca visiva viene utilizzata anche da numerose aziende per monitorare la presenza o frequenza con cui i loro loghi vengono visualizzati sui social network e in rete.

Le reti neurali CNN

Le reti neurali più usate nella computer vision sono le reti CNN (reti convoluzionali), che si ispirano al funzionamento della corteccia visiva primaria, ovvero di quella parte del cervello che elabora le informazioni visive e rende possibile agli uomini la visione.

Gli assistenti digitali 'spie'

Dietro gli assistenti vocali c'è una IA. Gli Assistenti vocali attuali sono ad esempio:

- Smartphone: Siri, Google Now/Assistant, Cortana
- Computer: Cortona
- Device: Alexa/Echo, Google Home/Nest, HomePod

Attivazione senza consenso. Ormai numerosi casi ci hanno fatto capire che questi assistenti vocali ascoltano ed elaborano quello che diciamo anche prima della famosa parola di attivazione: Hey Siri, Alexa/Echo, Ey Google. Quindi agiscono di fatto come delle spie[23]. A chi non è capitato di trovare pubblicità di quello di cui parliamo, e a volte appaiono anche a nostri familiari… e a volte incrociano i dati in una maniera "curiosa" tanto da sbalordire o inquietare.

[23] Smartphone "spioni": il Garante Privacy avvia un'istruttoria, Altalex 23/11/2021.

La Trasmissione dei dati

- Ci sono webcam che inviano dati a diversi server di quello dello streaming per l'utente.
- Ci sono SmartTv che inviano dati a piattaforme come Facebook e Netflix senza autorizzazione (anche se non avete messo il vostro account).
- I dati sullo Smartphone come l'archivio fotografico, la rubrica, il calendario, ecc. sono trasmissibili ad altri server quando viene accettato il consenso al solo accesso da parte delle App. I nostri dati audio vengono inviati a qualche server? Sembra di sì vedendo alcuni incroci a distanza.
- Alcune app per la registrazione del ciclo mestruale, rapporti sessuali se protetti o meno, masturbazione ecc. condividevano i loro dati a Facebook senza il consenso delle utenti.

Non vengono raccolte solo le informazioni personali ma possono anche riferirsi *a chiunque si trovi nello stesso ambiente* e riguardano[24], ad esempio:

- Scelte, preferenze e abitudini relative a stili di vita, consumi, interessi, ecc.;
- Caratteristiche biometriche, come ad esempio quelle della voce e del volto, se i dispositivi sono dotati di videocamera;
- Geolocalizzazione: posizione, percorsi abituali o frequenti, domicilio, indirizzo del posto di lavoro, ecc.
- Numero e caratteristiche (età, sesso, ecc.) delle persone che si trovano nell'ambiente in cui operano
- Stati emotivi

"IoT" – Internet of things

La particolarità dei dispositivi di essere connessi a Internet ("IoT" – Internet of things), permette che i dispositivi IoT possano dialogare amplificando la possibilità di raccolta, incrocio dei dati e diffusione di informazioni personali, che l'IA può elaborare a volte in modo sorprendente o preoccupante[25].

In diversi casi nelle policy degli social media o degli assistenti digitali si ammette elegantemente l'acquisizione di dati.

Lo Specchio biometrico

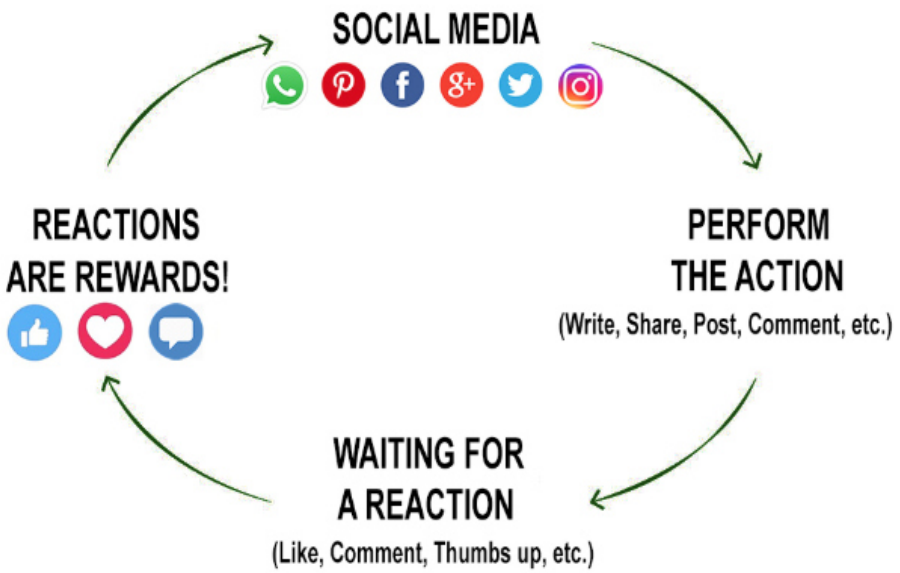
Molti social che ci invitano a farci foto e video, con una scansione del viso ottengono informazioni sull'utente, come:

- il grado di felicità,
- il senso di responsabilità
- la stabilità emotiva

Così si arriva a un'identificazione univoca. Questo metodo è utilizzato da TikTok come descritto nella sua privacy policy. Se il Social Media individua lo stato emotivo dell'utente, potrebbe suggerirgli contenuti che rinforzano questo stato. Il rischio è che l'identità di una persona si formi su ciò che vede su internet, diventando ciò che l'algoritmo vuole[26].

L'IA per lo stimolo di dipendenze digitali

[24] Tutti spiati?, Rai Presadiretta 10/02/2020.
[25] Assistenti digitali: i consigli del Garante per presidiare la privacy**,** Elisa Chizzola Il sole 24 ore 05/09/2023.
[26] La dittatura dell'algoritmo, Rai Presadiretta 20/03/2023.



L'IA nella cattura dell'utente

I social network hanno "bacheche" dove ci propongono in modo più o meno arbitrario contenuti e pubblicità. Queste bacheche sono il prodotto di IA in base ai dati raccolti dai nostri smartphone e dai nostri movimenti in rete. Più tengono gli utenti "incollati", più hanno dati disponibili che fanno aumentare, conseguentemente, le possibilità di guadagno.

Quando Facebook ha cambiato le notifiche da blu a rosso ha attirato molto più tempo gli utenti che sono stimolati di più a farle sparire.

La gestione della pubblicità

Gli strumenti basati sull'IA possono contribuire a esaminare centinaia o migliaia di varianti di target e di budget per le pubblicità, a identificare e segmentare il pubblico, a progettare creatività per le pubblicità, a testare le pubblicità e a perfezionare la velocità e l'efficienza in tempo reale per raggiungere i risultati migliori. I più deboli possono venirne suggestionati.

Lo scrolling e dopamina

È una tecnica che stimola i sistemi di ricompensa che generano dipendenza: la curiosità di un nuovo contenuto (social media) o di una nuova vincita (giochi di azzardo, videogiochi) Si chiama *Nucleus accumbens*, quella parte del cervello responsabile della brama e del desiderio. Lo scrolling abbinato all'elaborazione di bacheche gestite da IA può generare un loop senza fine. La media giornaliera dei nostri studenti su TikTok è di 3 ore e 2 su Instagram.

La gestione dei contenuti e rafforzamento delle proprie tendenze.

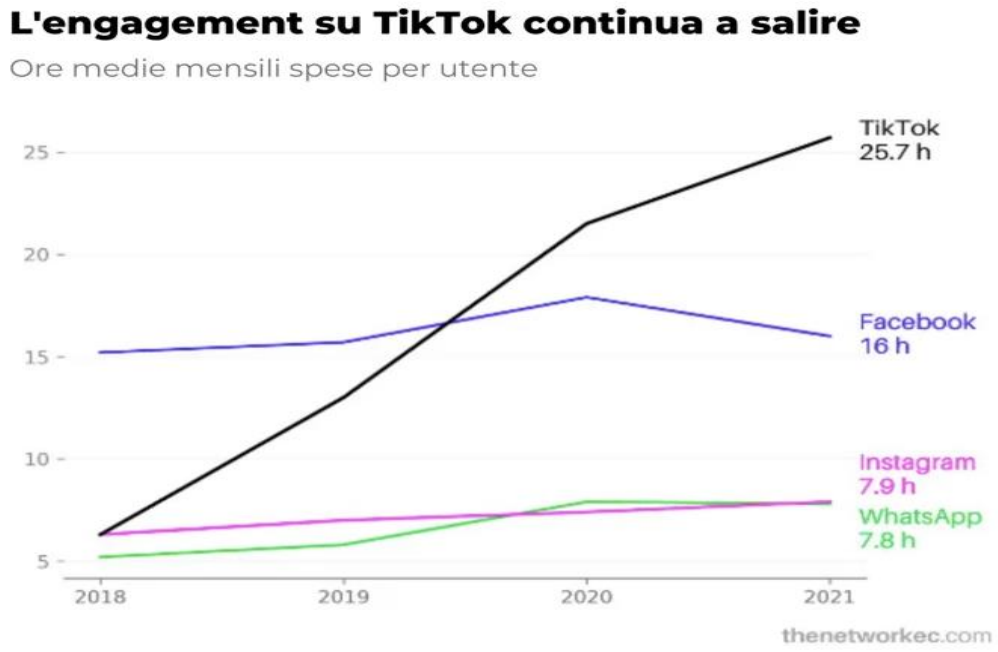
Tutto ciò che si pubblica viene quindi analizzato e confrontato con gli altri contenuti sulla piattaforma. Le IA tendono a rafforzare le proprie tendenze ma un adolescente in crisi può trovarci la morte:

- Il caso Molly, Londra 2017: sentenza storica per la morte di Molly: "i contenuti di Instagram e altri social responsabili del suicidio di una 14enne". Vittima dell'IA dei social (principalmente Instagram e Pinterest) che la spingeva all'autolesionismo, ansia e depressione. Dopo queste sentenze dalla spregiudicatezza del marketing qualche social

cerca di essere più vigile[27]. L'IA di Pinterest gli inviava pure e-mail con foto di lance insanguinate. Instagram video con gesti di autolesionismo, immagini con bambini col cappio al collo... Non bannava i commenti dei bulli. Instagram aveva cambiato l'IA della bacheca proprio l'anno prima (2016): da post in ordine cronologico a post mirati.

- TikTok: è facile simulare l'età dei 13 anni per entrare e anche qui si è verificato con un test che simulandosi un adolescente con disturbi alimentari proponesse video di anoressiche o gente con il sondino nasogastrico che facevano a gara a mostrarsi più sofferenti degli altri.
- Twitter oggi X per anni non aveva particolari o efficaci restrizioni riguardo alla pedopornografia. Non si era interessato più di tanto a limitarne la diffusione. Altri social network da segnalare sono Reddit, Snapchat e YouTube.

Social Network e sviluppo dopamina: Like or not Like. Gli elementi di rinforzo: like, reazioni, retweet, follower, condivisioni danno una scarica di dopamina detta anche molecola della ricompensa, che di rende dipendenti dalle cose che la provocano.
Justin Rosenstein, l'uomo che ha inventato il Like rinuncia a Facebook: "volevo regalare un po' di ottimismo invece è diventata una killer App". Ci si è resi conto, come nel caso di YouTube, il dislike generava depressione. O anche il confronto con il numero di like, come nel caso di Instagram. Le IA che poi fanno la loro parte, dove scelgono quali contenuti facilitare a diventare virali, spesse volte senza merito.
Accade spesso che gli esseri umani inventino con le migliori intenzioni per poi scoprire che qualcosa non va: "Tra poco potremmo essere gli unici in grado di raccontare com'era la vita prima che accadesse - immagina Rosenstein -prima, cioè, che Apple, Facebook, Google, Twitter, Instagram o Snapchat imparino a controllare le nostre menti, solo perché pensiamo che questo sia il traguardo da raggiungere"[28].



Frances Haugen ex dipendente di Facebook, oggi Meta, denuncia: ecco il Video.

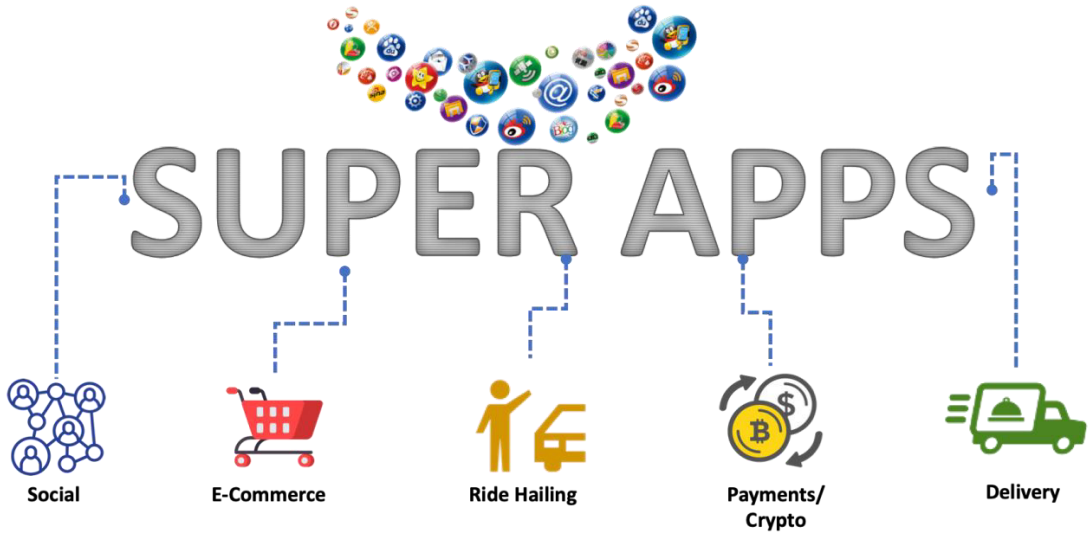
[27] La scatola nera, Rai Presadiretta 20/03/2023.
[28] D. Michielin *Così i social network plasmano il cervello,* Repubblica 04/05/2017; Justin Rosenstein, l'uomo che ha inventato il Like rinuncia a Facebook, di Gaia Scorza Barcellona, La Repubblica, 09/10/2017.

L'IA, Democrazia e Politica

Abbiamo notato come gli algoritmi di Meta e TikTok possano essere armi politiche usate per influenzare il cittadino medio che va a votare, l'esempio più lampante è il caso di facebook e lo scandalo Facebook-Cambridge Analytica all'inizio del 2018. Cambridge Analytica aveva raccolto i dati personali di 87 milioni di account Facebook senza il loro consenso e li aveva usati per scopi di propaganda politica.

SAI, SuperApp e supercontrollo

Definizione: Sono applicazioni mobili o web in grado di fornire molteplici servizi, tra cui l'elaborazione di pagamenti e transazioni finanziarie, diventando di fatto una piattaforma on-line di commercio e comunicazione onnicomprensiva che abbraccia molti aspetti di vita personale e commerciale.



Alias "cosa sono": sono applicazioni che vogliono coprire tutte le app in un'unica superapp. Un po' ricorda come la IA vuole diventare SAI.

Esempi notevoli di super-app

- WeChat: prima app in Cina. 1,32 miliardi di utenti la usano ogni mese (1,01 miliardi di questi sono cinesi)[29].
- Grab: di Singapore è la prima App nel sud-est asiatico, con 180 milioni di utenti. Qui, nel marzo 2018, Uber ha unito i suoi servizi con quelli di Grab in cambio di una quota di proprietà del 27,5% in Grab.
- Altre: c'è TataNeu in India, c'è l'indonesiana Gojek. Lark in USA, Giappone e Asia Pacifico; Alipay (in alcuni casi sono rivendicazioni personali, perché ancora non hanno funzionalità complete)
- Tentativi di diventare SuperApp in occidente: sono stati attribuiti a PayPal e a X (ex Twitter).

[29] 10 nuove statistiche Wechat nel 2024, Eddie Hurst, Ertahweb 20/07/2023.

Aspetti critici

- *Marketing:* Ci troviamo non solo di fronte a una questione tecnologica ma anche economica, dove in cambio di performance tecnologiche devi essere disposto a dare dati e riceverne, subendo una pressione psicologica funzionale al marketing.
- *Esclusività*: Le SuperApp hanno una pretesa morbosa verso l'utente "*devi essere tutto mio solo mio*".
- Controllo totale: Memorizzano quantità maggiori di dati personali, perché maggiormente coordinati rispetto a scambio di App di diversi produttori. Si corre un rischio maggiore di diventare vittime di gravi violazioni dei dati. Ciò è accaduto nel 2020 con Tokopedia[30] (fusa con Gojek del 2021 e comprata da TikTok per il 75% nel 2023), quando i dati di 91 milioni di utenti sono stati rubati e messi in vendita da hacker per 5000 euro.
- *Ricatto:* L'essere bannato da una superapp può mettere in seria difficoltà la vita delle persone con la perdita di accesso a servizi fisici e digitali.

In conclusione, ci chiediamo

1. Come vediamo l'essere osservati continuamente attraverso i nostri dispositivi?
2. E essere influenzati dai suggerimenti che ci vengono proposti dagli infiniti algoritmi adottati dai social?

[30] https://en.wikipedia.org/wiki/super-app

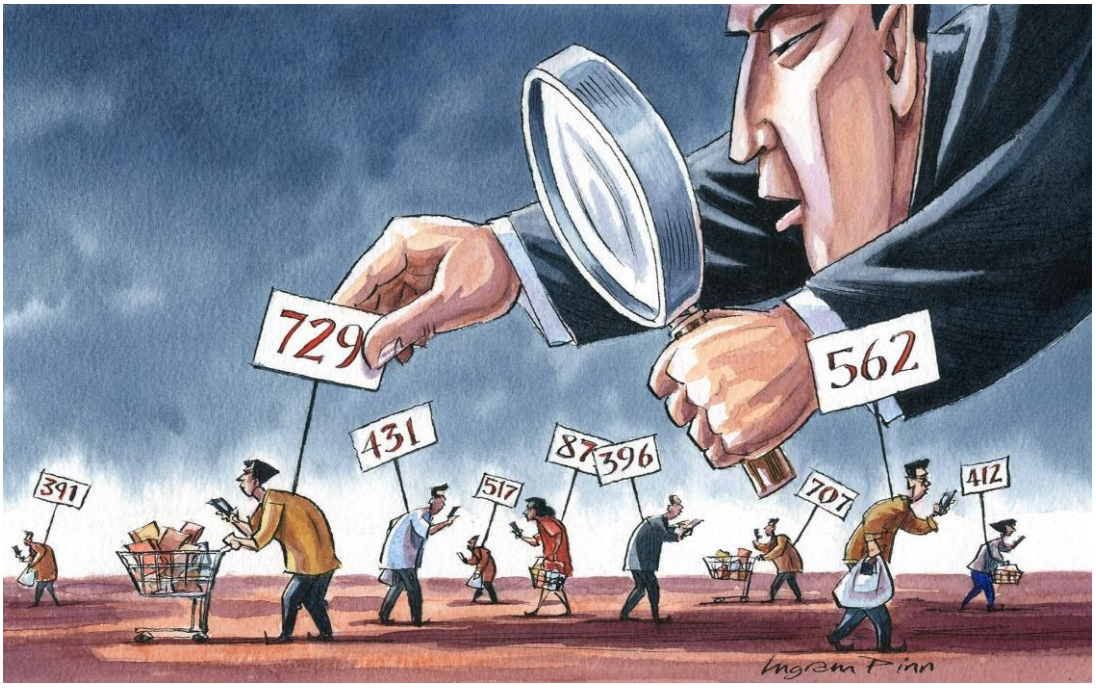
PARTE 2. I sistemI di credito sociale

Il sistema di credito sociale cinese

È un'iniziativa del governo per classificare con un punteggio persone e aziende sulla base del loro comportamento nella società.

Questo sistema si basa sulla sorveglianza di massa e sull'analisi dei big data, resa possibile dall'intelligenza artificiale. Il funzionamento è molto semplice: se una persona commette azioni buone e utili per la società, viene premiata con dei vantaggi e bonus sociali, viceversa per le persone che compiono atti deplorevoli, che ottengono svantaggi e restrizioni.

Nonostante il sistema sia ancora in fase di sperimentazione e il modo in cui è organizzato cambia nelle varie città e provincie, l'idea è quella di attribuire un punteggio numerico a persone e aziende sulla base delle loro azioni. Azioni buone = punteggio alto. Azioni cattive = punteggio basso.



Quali azioni vengono ritenute buone e quali cattive

Si considerano azioni positive: donare il sangue, fare una buona raccolta differenziata, fare la carità, lodare il governo sui social, finanziare le opere del governo, fare soffiate su persone che appartengono a culti eretici, pagare le tasse, fare volontariato…

Si considerano azioni negative: compiere atti criminali, non adempiere agli obblighi militari, protestare contro le autorità, appartenere a un culto eretico, non pagare una multa, fumare nei trasporti pubblici, diffondere fake news, attraversare fuori dalle strisce pedonali, non tenere il cane al guinzaglio.

Ricompense e punizioni

Le ricompense per le persone che hanno un punteggio elevato sono varie, e comprendono avere la precedenza negli ospedali e per quanto riguarda la sanità, la

possibilità di mandare i figli in una scuola facoltosa, la facilità con la quale si può accedere a mutui e finanziamenti generali, non pagare le cauzioni di veicoli come le bici.

Sono varie anche le punizioni per chi ha un punteggio basso del sistema, e includono l'impossibilità di accedere a un mutuo, di mandare i figli in una scuola facoltosa, di viaggiare con aerei e treni veloci e in generale si hanno una serie di svantaggi per quanto riguarda la sanità, e in alcuni casi, questa viene del tutto negata.

IA e i sistemi di credito per il controllo sociale

Il sistema è reso possibile dalla sofisticata tecnologia che possiede la Cina per la sorveglianza di massa. La Cina monitora in modo rigido e attivo le attività dei suoi cittadini online: i social, i forum online e altri servizi internet sono strettamente monitorati dal governo. Non è un caso che il paese del Dragone sia al 1° posto per l'implementazione e lo sviluppo di sistemi di riconoscimento facciale, reso possibile da un numero enorme di telecamere nel paese, concentrate in special modo nelle città (si stima che nel 2020 ne fossero presenti 600 milioni).

Per fare pressione sui cittadini affinché non compiano azioni negative, non manca l'umiliazione e la gogna mediatica per chi ha un punteggio basso nel sistema, che spesso viene inserito nelle cosiddette liste nere. Se già sono presenti cartelloni e siti web per vedere le persone nelle liste nere, in alcune città siamo arrivati al punto in cui se si chiama una persona nella lista nera interviene la segreteria telefonica per avvertire che si sta per parlare con una persona "poco affidabile".

Come valutano i cittadini i sistemi di credito sociale

Un unico studio è stato fatto per capire il tasso di gradimento dei cittadini sul sistema di credito sociale, ed è quello compiuto dalla Libera Università di Berlino nel 2018. In un sondaggio online proposto a 2109 persone cinesi, l'80% di questi approvava il sistema, con motivazione principale la riduzione della criminalità.

Obiettivo dichiarato del governo è quello di ristabilire una società armoniosa e civile, in cui, grazie al sistema, ci si può fidare o meno del prossimo.



Sistemi di valutazione sociale in Italia

Anche in Italia sono presenti degli esperimenti del sistema di credito molto limitati in confronto al caso cinese e su base volontaria: a Bologna, scaricando una app, c'è un sistema che premia i cittadini che fanno una buona raccolta differenziata e che utilizzano i trasporti pubblici. A Fidenza (provincia di Parma) è presente un meccanismo a punti per valutare il comportamento delle persone che vengono ospitati negli alloggi pubblici. A Bergamo se si va al lavoro in bici, si ottengono dei buoni di 25 centesimi per ogni km percorso.

Il recente documenti del Parlamento Europeo, AI Act, vieta nella Comunità Europea, i sistemi di credito sociale come quello cinese.

Valutazioni generali

Il rischio del sistema di credito sociale cinese è che si arrivi allo scenario dipinto nel libro 1984, in cui le persone sono continuamente spiate e non possiedono quindi una goccia di libertà. Se nel bestseller di George Orwell un bambino denunciava il padre alla polizia per aver criticato il Grande Fratello, non è tanto lontana la possibilità che i legami familiari si spezzino a causa del sistema di credito sociale, perché magari non si è portato il cane al guinzaglio o si attraversa fuori dalle strisce pedonali.

Le preoccupazioni riguardano la violazione della privacy, e il rischio principale è che si arrivi a un controllo totale sulla società, in cui ai cittadini è negata la libertà individuale e l'autonomia decisionale, con l'obbligo di conformarsi a determinati comportamenti per ottenere un punteggio elevato e accedere a bonus sociali. Perché sì, è importante disincentivare a compiere azioni negative e criminali, ma fino a che punto è accettabile che un governo eserciti un controllo così stretto e rigido sul comportamento dei suoi cittadini? Il rischio è quello di trasformare la società in una realtà in cui la libertà individuale viene sacrificata sull'altare della conformità sociale. Ma forse, ci sta bene così.

PARTE 3. IA e neurodiritti

La lettura del pensiero

All'inizio della storia delle neurotecnologie[31] possiamo collocare il pletismografo, un macchinario inventato nel 1875 dal medico torinese Angelo Mosso capace, a suo dire, di «scrivere anche quelle emozioni che non si dipingono sul volto o si rivelano troppo debolmente con i battiti del cuore e con l'affanno della respirazione»[32].

Il macchinario venne successivamente adottato in modo assai discutibile da Cesare Lombroso con l'intento di determinare la veridicità delle dichiarazioni degli indiziati[33].

Ma il fatto che solo in tempi relativamente recenti abbiamo cominciato a dotarci di strumenti tecnologici che non dovrebbero farci dimenticare *che il desiderio di leggere il pensiero altrui popola il nostro immaginario da sempre*.

Quotidianamente, facciamo esperienza di come sarebbe un grande vantaggio poter leggere nei pensieri altrui evitando i fraintendimenti che nascono dal non sapere cosa passa per la testa del nostro interlocutore e, al tempo stesso, intuiamo come questa possibilità potrebbe dare adito ad usi impropri qualora venisse usata contro la volontà dell'altro.

Allo stato attuale, siamo lontani anni luce dalla possibilità di realizzare compiutamente questa possibilità, ma in questi ultimi trent'anni le neuroscienze hanno compiuto passi da gigante nelle tecniche di lettura del pensiero e i moderni dispositivi ci permettono oggi di cogliere i segnali di attività cerebrale con una precisione senza precedenti. Ma ciò che ha dato un impulso decisivo alla ricerca è stata la possibilità di utilizzare l'IA per decodificare l'enorme quantità di dati proveniente dai dispositivi.

La possibilità di decifrare direttamente i segnali cerebrali da un lato e la possibilità di collegarli alle macchine dall'altro ha permesso ai pazienti tetraplegici di utilizzare dei comuni tablet, scrivere e-mail, chattare, fare streaming musicale o guardare video su YouTube esclusivamente tramite l'attività cerebrale. Ma i progressi delle interfacce cervello-computer (BCI, brain computer Interface) possono servire non solo in campo medico, ma anche in altri settori come il gaming, l'istruzione o il marketing[34].

---

[31] Quando parliamo di neurotecnologie intendiamo la dimensione applicativa delle neuroscienze ovvero quell'insieme di metodi e dispositivi che si interfacciano con il sistema nervoso per monitorarne e modularne l'attività.

[32] A. Mosso, *Sopra un nuovo metodo per scrivere i movimenti dei vasi sanguigni nell'uomo*, in Atti della Reale Accademia delle Scienze di Torino, IX, 1875, 37

[33] C. Lombroso, *La polizia scientifica*, in Archivio di psichiatria, n. 7 del 1886, 611-612.

[34] S. K. Mudgal, S. K. Sharma, J. Chaturvedi, A. Sharma, *Brain Computer Interface Advancement in Neurosciences: Applications and Issues, in Interdisciplinary Neurosurgery*, n. 20 del 2020, 100694.

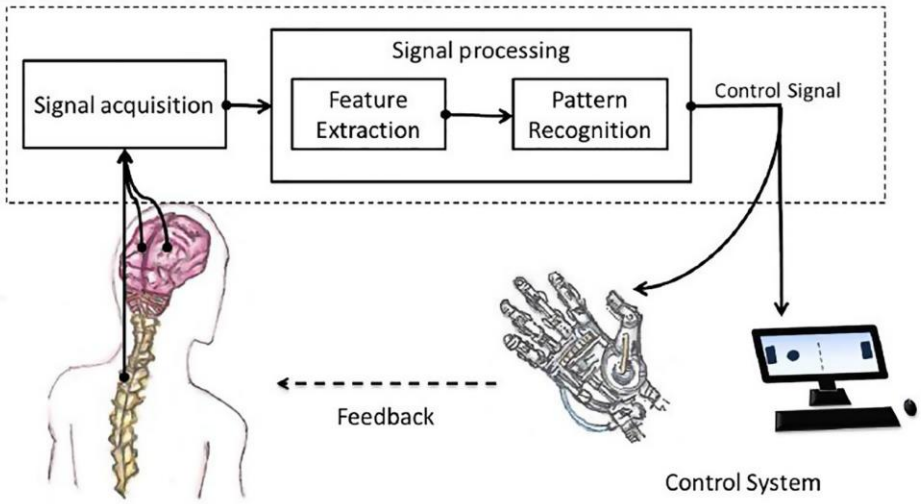


Fig. 2. Components of BCI in an illustration of transmission of signal between input and output with a sequence of processing steps in between. Signals are acquired by electrodes and then translated into a control signal for a machine-like neuro prosthesis or wheelchair depending on its intentional use [23].

Le tre principali modalità attraverso cui convergono intelligenza naturale umana e intelligenza artificiale sono:

1. l'intelligenza artificiale usata per decodificare l'intelligenza naturale.
2. l'integrazione di intelligenza naturale umana e intelligenza artificiale per controllare macchine digitali.
3. l'uso dell'intelligenza artificiale per aumentare o potenziare l'intelligenza naturale umana.

In questa esposizione analizzeremo alcune tecnologie neurali con risvolti nel campo medico, nello specifico in quello del miglioramento della qualità della vita in persone affette da paralisi totali, in stato vegetativo o simili.

Ci siamo fatti alcune domande:

1. È legittimo leggere il pensiero delle persone, se fosse possibile, senza la loro autorizzazione?
2. Se potessimo cancellare i ricordi delle persone, sarebbe legittimo?
3. Siamo in grado di manipolare la mente per far pensare agli altri quello che vogliamo?

Tecnologie di lettura del pensiero

Le moderne BCI possono essere distinte in due tipi: invasive e non invasive.

Le BCI invasive registrano la segnalazione cerebrale attraverso l'impianto chirurgico di serie di elettrodi collegati al sistema nervoso centrale.

Le BCI non invasive interfacciano la segnalazione cerebrale tramite tecnologie di *neuroimaging*.

Scanner cerebrale non invasivi

- *Risonanza magnetica funzionale (fMRI).* Attualmente lo scanner cerebrale più avanzato di cui disponiamo è il cosiddetto fMRI (functional Magnetic Resonance Imaging). Questo strumento mappa il flusso sanguigno nel cervello e restituisce all'operatore medico una fotografia dello stesso. Analizzando le parti del cervello che si attivano in determinati momenti si possono ricostruire delle immagini e/o delle frasi.
- L'elettroencefalografia (EEG). Alternativa alla risonanza magnetica e in grado di rilevare i segnali con una maggiore rapidità, l'elettroencefalografia (EEG) scansiona l'attività

neuronale grazie ad elettrodi applicati al cranio. Il limite di questa tecnica risiede nel fatto che gli elettrodi rilevano gli impulsi di interi gruppi di neuroni e conseguentemente può determinare con una precisione di pochi centimetri appena la regione cerebrale da cui scaturisce l'attività.

- Tomografia a emissione di positroni (PET). La PET rileva coppie di raggi gamma emessi indirettamente da un radioligando (una sostanza biochimica radioattiva) in grado di emettere positroni (le antiparticelle degli elettroni), introdotto nel corpo attraverso una molecola biologicamente attiva chiamata "tracciante radioattivo". L'applicazione di questa tecnica allo studio del cervello si basa sul presupposto che le aree a maggiore radioattività siano associate a una maggiore attività cerebrale. Come nel caso della fMRI, anche nella PET quello che viene misurato indirettamente è il flusso di sangue verso le diverse parti del cervello, il quale è, in generale, ritenuto correlato con l'attività neuronale.
- La magnetoencefalografia (MEG). La MEG è una tecnica di neuroimmagine funzionale per la mappatura dell'attività cerebrale mediante la registrazione dei campi magnetici prodotti dalle correnti elettriche che si verificano naturalmente nel cervello.

Cuffia con lettore di onde cerebrali. Questa cuffia è un oggetto sviluppato in Australia, dai ricercatori del GrapheneX-UTS Human-centric Artificial Intelligence Centre della University of Technology Sydney (UTS). Consiste in un sistema non intrusivo e portatile che, sfruttando un EEG (ElettroEncefaloGramma), riesce a decodificare i pensieri e trasformarli in testo. Allo stato attuale questo macchinario riesce a convertire in pensieri con una precisione di circa il 40% sulla scala BLEU-1 (Il punteggio BLEU è un numero compreso tra zero e uno che misura la somiglianza del testo tradotto automaticamente con un insieme di traduzioni di riferimento di alta qualità). Per avere un riferimento, i traduttori e gli assistenti vocali dei telefoni si piazzano su una media di circa 90% sulla medesima scala.

Scanner cerebrale invasivi: il chip nel cervello

L'azienda Neuralink ha sviluppato un prodotto che, al contrario dei precedenti, è invasivo in quanto viene impiantato all'interno del corpo umano. Questa misura è necessaria visto che esso si avvale di 1024 elettrodi connessi al sistema nervoso centrale per sondare gli impulsi elettrici che fungono da messaggeri nel nostro sistema nervoso. Intercettando questi ultimi è possibile decodificare i segnali inviati dal cervello e quindi, dopo una lunga fase di calibrazione specifica al soggetto in cui è impiantato il chip, tradurre i dati acquisiti in testo o movimenti di un mouse.

Il ruolo chiave dell'IA nelle nanotecnologie e provocazioni al Diritto

Entrambe le tipologie di BCI stabiliscono un'interazione diretta tra il cervello dell'utente e un dispositivo informatico attraverso un ciclo a quattro fasi.

La prima fase riguarda l'input, cioè la generazione di attività cerebrale specifica da parte dell'utente in risposta a uno stimolo.

La seconda fase riguarda la misurazione e la registrazione dell'attività cerebrale.

Nella terza fase, i dati neurali grezzi misurati nella seconda fase devono essere decodificati nelle loro caratteristiche principali e susseguentemente classificati. Questa elaborazione permette di estrarre le caratteristiche rilevanti dal segnale. In questo processo di decodifica e classificazione, l'intelligenza artificiale ha un ruolo fondamentale. Specialmente l'apprendimento automatico. Gli algoritmi, infatti, vengono addestrati al fine di decodificare correttamente i segnali neurali registrati e amplificati, nonché a classificare correttamente le derivanti categorie di dati.

L'IA ha permesso di superare il limite contro cui si infrangevano i neuroscienziati che di fronte ad una scansione cerebrale non erano in grado di dire cosa qualcuno stesse vedendo, sentendo o pensando nello scanner.

Grazie alla potenza predittiva dei modelli linguistici di intelligenza artificiale, siamo grado di riprodurre, con un livello di precisione sorprendente, le storie che una persona ha ascoltato o immaginato di raccontare nello scanner. Il decodificatore può anche indovinare la trama di un cortometraggio che una persona ha guardato nello scanner, anche se con minore precisione.

"Quello che otteniamo è ancora una sorta di 'succo' o, meglio, una parafrasi, di quella che era la storia originale", dice Alexander Huth, neuroscienziato computazionale all'Università del Texas ad Austin e autore senior dello studio.

Nell'articolo troviamo alcune trascrizioni di ciò che i partecipanti allo studio hanno ascoltato: "Mi sono alzato dal materasso ad aria e ho premuto il viso contro il vetro della finestra della camera da letto aspettandomi di vedere degli occhi che mi fissavano e invece ho trovato solo il buio".

Esaminando le scansioni cerebrali della persona, questo è ciò che è stato decodificato dalla macchina: "Ho continuato ad avvicinarmi alla finestra e ad aprire il vetro, mi sono alzato sulle punte dei piedi e ho sbirciato fuori, non ho visto nulla e ho alzato di nuovo lo sguardo, ma non ho visto nulla"[35].

Senza l'IA a svolgere questo passaggio si impiegherebbero mesi per produrre una singola frase. Dando invece i dati in pasto all'IA si ottimizzano i tempi anche se la precisione non è delle migliori. Questo perché si devono fare delle associazioni e visto che i dati forniti sono poco precisi, per svariati motivi differenti in ciascun caso, queste trascrizioni pensiero-testo/azione sono carenti. Tramite una calibrazione (sessione in cui lo stesso soggetto utilizza l'apparecchio per un quantitativo di ore considerevoli) si possono mappare le interazioni più semplici, ma per quelle più complesse è difficile trovare una chiave di lettura. Questo è dovuto anche al fatto che ogni cervello è diverso dall'altro quindi, ogni volta che si cambia soggetto bisogna iniziare nuovamente da zero.

Una volta decodificati, i segnali possono essere tradotti in output che implicano solitamente l'esecuzione dell'azione inizialmente intesa o desiderata o ritenuta benefica per l'utente attraverso il controllo delle applicazioni interfacciate dal BCI.

Livello di avanzamento della ricerca e questioni legali

Negli ultimi decenni la neurotecnologia e l'intelligenza artificiale hanno reso leggibili i processi dell'intelligenza umana e nel prossimo futuro la combinazione di intelligenza naturale e artificiale entrerà sempre più a far parte delle nostre vite.

Ma la possibilità di accedere ad una immensa mole di dati celebrali (anche detti dati neurali) solleva diverse domande sull'intimità e la riservatezza dell'attività cognitiva (neurale, cerebrale, psichica, ecc.) e altre più ampie questioni etiche, che presentano ricadute anche sul piano del diritto.

Attualmente si sente parlare di lettura del pensiero e violazione della privacy nel modo più estremo possibile, questo però va preso con le pinze in quanto la ricerca è ben lontana dal poter affermare di leggere il pensiero di una persona o, addirittura, di modificarlo e influenzarlo. Detto questo è anche giusto guardare con occhio critico il futuro, sia quello più vicino sia quello lontano dal nostro tempo, e iniziare a gettare le basi per uno sviluppo regolamentato di questi apparecchi nella loro interezza. forse

La scienziata Fei-Fei Li sostiene che non vi è niente di artificiale nell'intelligenza artificiale: «è ispirata da persone, è creata da persone e, cosa più importante, incide sulle persone. È uno strumento potente che si sta solo ora cominciando a capire come funzioni, e questa è una profonda responsabilità»[36].

[35] TANG J. e altri, *Semantic Reconstruction of Continuous Language from Non-Invasive Brain Recordings,* in «Nature Neuroscience», Vol. 26, pp. 858-66, 2023.

[36] Fei-Fei Li: https://www.wired.com/story/fei-fei- li-artificial-intelligence-humanity/

Nuovi diritti necessari: la nascita dei Neurodiritti

Seguendo il discorso iniziato nel punto precedente sarebbe interessante riflettere su questi quattro punti:

1. Privacy mentale
2. Continuità psicologica
3. Integrità mentale
4. Libertà cognitiva

Diritto alla privacy mentale

Mira a proteggere gli individui da intrusioni non desiderate nei loro pensieri e processi mentali. Questo diritto dovrebbe consentire alle persone di proteggere le informazioni neurali, specialmente quelle elaborate al di sotto del livello di percezione cosciente. Le persone potrebbero essere particolarmente vulnerabili alle violazioni della privacy mentale poiché hanno un limitato controllo volontario sui segnali cerebrali. Con il progresso tecnologico, c'è il rischio di una corsa agli algoritmi che elaborano dati neurali per creare modelli predittivi del comportamento e delle preferenze dei consumatori. Ciò potrebbe portare alla micro-targhettizzazione e alla profilazione basata sulle caratteristiche neurocognitive degli individui.

Diritto alla continuità psicologica

Questo diritto mira a proteggere l'identità personale e la continuità della vita mentale dalle alterazioni esterne non volute. Si considerano inclusi in questo diritto interventi non autorizzati di modifica della personalità che non causano danni fisici o psicologici diretti, come ad esempio la neuromodulazione per scopi militari. Questo principio si applica anche a strategie di marketing invasive progettate per alterare preferenze o comportamenti.

Integrità mentale

Questo diritto, già riconosciuto dal diritto internazionale, dovrebbe essere esteso per proteggere dalle manipolazioni illecite e dannose dell'attività mentale abilitate dalle neurotecnologie. Queste minacce includono la neurostimolazione indesiderata, il neurohacking malevolo e la manipolazione potenzialmente dannosa della memoria. Si considera importante nel contesto della sicurezza nazionale, dove potrebbero essere giustificati interventi potenzialmente dannosi per raggiungere obiettivi strategici.

Libertà cognitiva

Questo diritto mira a proteggere la libertà degli individui di prendere decisioni autonome sull'uso delle BCI e di altre neurotecnologie. Gli individui adulti competenti dovrebbero essere liberi di utilizzare le BCI sia per scopi clinici che di potenziamento cognitivo, purché non violino le libertà altrui. È importante anche il diritto di rifiutare applicazioni coercitive, inclusi quelli implicitamente coercitivi, per prevenire la manipolazione forzata degli stati mentali dei cittadini da parte dello Stato, delle grandi aziende o attori malevoli.

Casi di cronaca e neurodiritti

Gli avanzamenti nel campo degli impianti neurotecnologici hanno raggiunto livelli straordinari, ma è l'integrazione dei sensori cerebrali nella tecnologia comune che avrà un impatto significativo sulle nostre vite.

I casi cinese e australiano sollevano sfide sul piano etico che investono l'accesso ai dati personali di natura intima e mentale.

Monitorare la stanchezza sul lavoro

Nel 2018 il *South China Morning Post* riferiva che decine di fabbriche e aziende avevano cominciato a usare "dispositivi di sorveglianza cerebrale" per monitorare lo stato di salute mentale dei loro dipendenti.

Tramite dispositivi collocati all'interno di cappelli e caschi, si intendeva svolgere un'analisi costante delle onde cerebrali dei lavoratori per valutare in tempo reale lo stato di salute mentale dei dipendenti, in modo da intervenire rapidamente nel caso si riscontrasse un elevato carico di stress, di ansia, paura o rabbia.

La stabilità mentale di un pilota d'aereo o di un treno ad alta velocità, ma anche di un responsabile della fornitura energetica di un'intera regione oppure quella di un comandante di un plotone d'esercito venivano descritti come esempi in cui tale tecnologia avrebbe potuto essere impiegata con effetti positivi.

Al di là del fatto che molti scienziati accolsero con scetticismo l'annuncio sottolineando la scarsa affidabilità dei dati rilevati attraverso l'elettroencefalogramma realizzato da quel genere di dispositivo, quel che preoccupava le opinioni pubbliche internazionale era il fatto che un tale esperimento venisse condotto all'interno di una nazione guidata da un governo autoritario.

Ma tre anni prima, nel 2015, sulle pagine del quotidiano australiano *Australian Financial Review* era stata pubblicata un'inchiesta in cui si parlava del monitoraggio mentale operato sui lavoratori di imprese australiane. Nell'articolo si citava lo "SmartCap", una fascia cerebrale dorata di sensori EEG che forniva un punteggio di valutazione del rischio basato sulla resistenza al sonno di chi indossava il dispositivo.

Nonostante i limiti significativi della tecnologia e la complessità computazionale implicita nel fatto di classificare i segnali delle onde cerebrali come «assonnati», si riuscivano ad ottenere risultati efficaci tanto quanto il riconoscimento facciale e le telecamere interne.

Ma, a differenza del caso cinese, la preoccupazione per la *privacy* mentale finiva per essere messa in secondo piano rispetto all'interesse molto concreto di sapere se l'autista di un camion da quaranta tonnellate che sfrecciava in autostrada tenesse gli occhi bene aperti.

Tenere sotto controllo l'attenzione degli studenti…

L'anno successivo, il 2019, divennero virali la foto di alcuni bambini di un istituto di eccellenza cinese, ordinatamente seduti sui banchi, mentre indossano una fascia nera intorno alla testa. Si trattava di uno strumento di rilevazione dotato di sensori per leggere le onde cerebrali e rilevare, in tempo reale, il livello di concentrazione degli studenti.

I rappresentati del governo cinese che aveva avviato la sperimentazione spiegavano che l'obiettivo della fascia era di offrire all'insegnante un *feeedback* immediato sull'efficacia della propria lezione ma furono in molti a pensare che più che esprimere un giudizio sul professore, rischiava di diventare uno strumento per valutare gli studenti.

Infatti, nell'articolo del *Wall street journal* che riportava la fotografia si raccontava che i dati degli studenti venivano comunicati, in tempo reale, non solo agli insegnanti ma anche ai genitori e allo stato.

Quello stesso anno, Emotiv presentava delle cuffie di sua invenzione che potevano essere usate per ascoltare musica o partecipare a conferenze ma, al tempo stesso, erano in grado di monitorare le funzioni emotive e cognitive di chi le indossava grazie a due elettrodi inseriti al loro interno.

Si trattava di dispositivi nati per monitorare e migliorare le prestazioni mentali ma anche in questo caso le preoccupazioni sollevate dal caso cinese venivano scalzate dalle

considerazioni sul fatto che sapere perché e quando subentra la distrazione poteva aiutare ad acquisire abitudini lavorative più sane.

Controllare la propria attività cerebrale e salute

I ricercatori della Ben-Gurion University del Negev, in Israele, hanno già sviluppato un dispositivo EEG indossabile detto Epiness che, sostengono, è in grado di prevedere gli attacchi epilettici anche un'ora prima che si verifichino, inviando un allarme sullo smartphone.

Grazie a questo dispositivo una persona epilettica potrebbe andare al lavoro in macchina senza il timore di fare un incidente lungo la strada oppure un paziente farmaco-resistente potrebbe assumere le benzodiazepine solo al bisogno, per alleviare un attacco imminente, evitando di sviluppare la tolleranza al farmaco.

In tutto il mondo, gli organi di controllo tendono a classificare Epiness e altri strumenti analoghi come dispositivi medici, sottoposti a regole stringenti sia prima sia dopo la vendita. Ma nel caso in cui questi dispositivi siano davvero in grado di fare ciò che proclamano i loro produttori, dovrà essere un esperto a decidere se e come potremo usarli?

Di fronte alla possibilità di offerta da questi dispositivi di accedere direttamente ai dati cerebrali, sono in molti a ritenere che i consumatori non ne sappiano abbastanza per interpretare correttamente i dati sul proprio cervello e sul proprio corpo, e che dunque siano destinati a compiere scelte sbagliate.

Esempi di Neurodiritti già introdotti in alcune Costituzioni e Leggi

Il precursore di tutto ciò è, inaspettatamente, il Cile che, già nel 2021, ha introdotto una sezione nella sua costituzione che regola e delimita l'area di lavoro per le neurotecnologie. L'idea di fondo è quella di paragonare i pensieri e i dati prodotti dal cervello ad un organo, rendendone illegale il traffico, la compravendita e la manipolazione. Un altro caso analogo è quello del Messico, che invece ha aggiunto due capitoli alla sua Carta dei Diritti Digitali del Cittadino. Il primo dei due contiene cinque neurodiritti mentre il secondo regola l'utilizzo dell'IA su una base etica.

Come i paesi e le varie parti del mondo stanno affrontando la questione dei neurodiritti e la regolamentazione.

EUROPA

- Unione Europea (UE): Approva il "AI Act" che classifica le IA in base al rischio e stabilisce procedure per la loro commercializzazione.
- Regno Unito e Svizzera: Optano per non sviluppare regolamentazioni specifiche sull'IA, ma applicare leggi esistenti in contesto.

NORD AMERICA

- Stati Uniti: Approccio liberale, senza restrizioni stringenti sull'IA, ma con direttive federali vaghe.
- Canada: Lavora su leggi che regolano IA ad alto rischio, con focus su sicurezza, diritti umani e applicazioni etiche.

SUD AMERICA

- Brasile: Sta sviluppando una legge per IA ad alto rischio, con organo di controllo e responsabilità civile.
- Cile: Introduce regolamentazioni limitanti sull'IA e neurotecnologie.
- Messico: Aggiunge sezioni riguardanti IA e neurodiritti alla Carta dei Diritti Digitali.

ASIA ED AUSTRALIA

- Cina: Ha leggi avanzate e organi di supervisione per gestire le IA.
- Giappone: Non possiede leggi esclusivamente inerenti all'IA, ma ha una solida struttura di leggi specifiche nei vari settori, le quali arginano e limitano già gli sviluppatori e gli utenti sotto molti punti di vista commerciali ed economici.
- India: Sta sviluppando una regolamentazione per IA ad alto rischio e ha una commissione per gli aspetti etici, sociali e legali delle IA.
- Australia: Adotta un approccio simile a Regno Unito e Svizzera, adattando leggi esistenti all'IA.

Conclusioni

Le neurotecnologie sono senz'altro qualcosa che migliorerà la vita di molte persone ma se si ipotizza un utilizzo improprio di esse si prospettano scenari sgradevoli, simili sotto molti aspetti ad alcuni universi fittizi reperibili nella letteratura internazionale. Urge perciò una delimitazione delle aree di lavoro che possono essere esplorate, ed aree in cui non ci si dovrebbe addentrare minimamente. Deve essere una decisione netta che lasci meno aree grigie possibili per impedire lo sviluppo di tecnologie oppressive per la razza umana.

- *Anche se oggi queste tecniche sono relativamente rudimentali, dove arriveremo tra 5 o 10 anni?*
- *Alcuni di noi ritengono che nella gamma di possibilità ragionevoli rientrino cose che sono, se non spaventose, abbastanza distopiche, e quindi pensano che sia il momento di cominciare a rifletterci.*

EVOLUZIONI DELL'IA E RISVOLTI ETICI

Liceo Scientifico e Classico G.P. Vieusseux Imperia *Classe 5A Classico*
Docente: Prof.ssa Raffaella De Meo
Relatore: Giulio Rossi

Il lavoro svolto mira a riflettere sul percorso evolutivo dell'IA e sui suoi risvolti etici. Analizzeremo la definizione dell'IA e la sua storia, gli aspetti positivi e negativi e le sue molteplici applicazioni in medicina e in altri campi. L'integrazione dell'intelligenza artificiale (IA) in medicina solleva diverse sfide etiche.

Sebbene da un lato l'IA offra innovazione e miglioramenti nell'assistenza sanitaria, dall'altro richiede che sia a servizio dell'uomo. L' approccio deve essere riflessivo ed innovativo, sinergico, interattivo tra medico e tecnologia, dove la tecnologia venga utilizzata dall'uomo per l'UOMO.

È necessario adoperarsi per un'IA a servizio dell'uomo, per un nuovo **UMANESIMO** in cui la tecnologia segua un **codice etico** a favore del progresso e del miglioramento della vita umana, senza discriminazioni in un'ottica globale. Risulta essenziale costituire, attraverso l'educazione, uno stadio fondamentale di consapevolezza sia per i fruitori dell'IA, che per gli innovatori e programmatori sui principi, sui benefici, sui rischi e sviluppi futuri.

Definire l'IA è una sfida complessa, comprende sistemi che imitano l'intelligenza umana per svolgere compiti vari. L'acquisizione e la gestione dei dati risulta fondamentale. Nel corso del tempo sono occorsi numerosi progressi di questa tecnologia, che oggi trova numerose applicazioni in ambito: medico-sanitario, ecologico, climatico, di ricerca, di applicazione tecnologica. Tuttavia emergono dubbi sull'affidabilità, la responsabilità e l'informazione ai pazienti sull'uso dell'IA. La collaborazione tra professionisti medici, esperti di etica e programmatori è cruciale per massimizzare i benefici dell'IA e mitigare i rischi etici.

È fondamentale il ruolo dell'**Europa** con la legge approvata il 13/03/24, che regolamenta per gli stati membri le applicazioni dell'IA e stabilisce l'IA office europeo come sostegno ai cittadini alle imprese e la valutazione di impatto dell'IA.

La **scuola** promuove la sensibilizzazione degli studenti nei confronti di questa nuova tecnologia per un uso: corretto, consapevole e responsabile.

Evoluzioni dell'IA e risvolti etici

Parlare dell'intelligenza artificiale è cruciale per comprendere il suo impatto trasformativo nella società moderna. Questa tecnologia sta rivoluzionando settori come la sanità, l'automazione industriale, l'educazione e molto altro. Discutere dell'intelligenza artificiale ci consente di esplorare le sue implicazioni etiche, come la privacy e la sicurezza dei dati. Inoltre, stimola la riflessione sull'equità e sull'accessibilità, assicurando che i benefici dell'IA siano equamente distribuiti. È fondamentale coinvolgere un pubblico diversificato in queste conversazioni, dalla comunità scientifica, ai responsabili politici, ai cittadini. Solo attraverso un dialogo aperto possiamo plasmare responsabilmente il futuro dell'intelligenza artificiale per il bene comune. Oggi più che mai è doveroso interessarsi al tema dell'IA, essendosi chiaramente delineato uno degli aspetti fondamentali di essa: l'IA costituirà infatti una vera e propria rivoluzione, volenti o nolenti, in ogni campo, arrivando a interessare ogni ambito della nostra vita. Per comprendere a pieno il fenomeno dovremmo cercare di definire in maniera chiara e quanto più esaustiva possibile il fenomeno dell'IA.

Definire **l'intelligenza artificiale** è una sfida ardua, poiché essa costituisce ad oggi un fenomeno talmente complesso e ramificato per cui darne una definizione univoca e globale è assai difficile. Si può dire che l'intelligenza artificiale sia un prodotto informatico, o una macchina in grado sia di interagire e di apprendere dall'ambiente esterno, sia di assumere decisioni con crescenti gradi di autonomia. L'intelligenza artificiale (IA) è dunque un ramo

dell'informatica che mira a creare sistemi in grado di eseguire compiti che solitamente richiedono l'intelligenza umana. Questi sistemi possono comprendere il linguaggio umano, possono apprendere, ragionare, risolvere problemi e adattarsi a nuove situazioni. Ma come si è evoluta l'IA nel corso del tempo?

La storia dell'intelligenza artificiale ha radici profonde, ma ha visto significativi progressi nel XX secolo. Uno dei momenti fondamentali è stato nel 1950, quando Alan Turing propose il "Test di Turing", un criterio per determinare se una macchina può essere considerate intelligente. Questo ha gettato le basi per lo sviluppo dell'IA. Negli anni '50 e '60, gli scienziati hanno sperimentato algoritmi e modelli matematici per simulare il pensiero umano. Tuttavia, il progresso è stato lento a causa delle limitate capacità di calcolo dei computer dell'epoca. Negli anni '80 e '90, l'avvento di computer più potenti e l'accesso a grandi quantità di dati hanno portato a una rinascita dell'IA. Gli algoritmi di apprendimento automatic hanno cominciato a mostrare risultati promettenti in campi come il riconoscimento vocale, la traduzione automatica e la guida autonoma. Negli ultimi decenni, l'IA ha fatto passi da gigante grazie alla combinazione di potenti algoritmi di apprendimento automatico, enormi quantità di dati e potenza di calcolo sempre maggiore. Le reti neurali profonde, in particolare, hanno rivoluzionato molti settori, dall'assistenza sanitaria alla finanza, dall'industria automobilistica alla pubblicità online. Oggi l'intelligenza artificiale sta pervadendo ogni aspetto della nostra vita e continuerà a farlo sempre di più, volenti o nolenti. La diffusione dell'intelligenza artificiale infatti è un fenomeno involontario ossia la volontà dei singoli individui non può determinare che esso avvenga o meno. Oggi si sta diffondendo principalmente un tipo particolare di intelligenza artificiale, il deep learning, alla base, per esempio, di chat GPT. Questa tecnologia è impiegata in innumerevoli ambiti, talvolta controversi. È applicata in campo bellico mostrando tutti i suoi limiti, poiché ha avuto spesso ricadute terribili anche sui civili. Apporta enormi benefici nel mondo dell'istruzione, in particolare agli alunni con bisogni educativi speciali o al cui insegnamento si interpongono barriere linguistiche. Viene spesso impiegata nella videosorveglianza pubblica, permettendo di riconoscere automaticamente volti e veicoli e di seguirne i movimenti, con tuttavia numerosi problemi di privacy per gli individui.

In seguito ci concentreremo sulle **applicazioni in medicina dell'IA**. Esse sono infatti numerose, ma pongono anche delle problematiche, soprattutto etiche, di cui è doveroso parlare. L'IA ormai è applicata in ogni campo: dalla diagnostica, alla formulazione e gestione delle terapie, all'assistenza dei pazienti, alla chirurgia, alla sistematizzazione dei dati relativi al paziente, all'analisi e condivisione dei big data derivanti da studi e trial clinici, fino alla scoperta di molecole potenzialmente utili in ambito farmacologico.

L'IA permette anche un rapido sviluppo di quella che ormai è nota con il nome di **telemedicina**, che trova applicazione soprattutto in ambito chirurgico. Può inoltre aiutare la classificazione dei casi, può suggerire le terapie migliori, analizzando quelle che si sono dimostrate più efficaci in casi analoghi a quello preso in esame. Anche in campo epidemiologico può aiutare a sviluppare modelli di previsione di diffusione delle malattie infettive, nonché aiutare nell'elaborazione di vaccini.

Particolare rilievo sta assumendo recentemente il **sistema Watson**, sviluppato da IBM. Esso è un sistema computerizzato in grado di rispondere a domande espresse oralmente attingendo ad un'enorme mole di dati. Dunque può assistere il medico attingendo da linee guida, studi clinici, pubblicazioni su riviste scientifiche, e molto altro. L'intelligenza artificiale è stata inoltre largamente impiegata durante la pandemia Covid-19 per contrastare le fake news, rilevandole e segnalandole. La possibilità di comparare un caso particolare ad una quantità sconfinata di casi già analizzati ha notevole applicazione nella **diagnostica per immagini**, quindi nell'istopatologia e nella radiologia. L'IA ha infatti la capacità di comparare l'immagine a casi simili, fornendo suggerimenti diagnostici allo specialista. L'intellige za artificiale ha permesso inoltre la sintesi e la scoperta delle proprietà antibatteriche dell'**Halicina**. Il primo antibiotico scoperto con IA (Machine learning) capace di essere attivo

sui ceppi resistenti ai comuni antibiotici ed impedirne il **polimorfismo.** Inoltre sta guidando lo sviluppo del trial clinico su **BPM 31510**, un farmaco di cui è iniziata la sperimentazione come antitumorale per i tumori solidi, quali in particolare i tumori del pancreas, I glioblastomi, liposarcomi mixoidi e fibrosarcomi. Altro notevole progresso tecnologico legato all'IA è il **robot chirurgico "Da Vinci",** che permette di operare a distanza, ed inoltre garantisce una serie di vantaggi per il recupero del paziente, in quanto la macchina è molto precisa ed affidabile per quel che riguarda le microincisioni, riducendo il sanguinamento operatorio e quindi la necessità di trasfusioni, e la riduzione del tempo di recupero per il paziente. Il Da Vinci è l'esempio più avanzato di questa tecnologia, anche se altri modelli analoghi sono già in via di sviluppo.

Tuttavia si riscontrano numerosi problemi: se sia possibile e conveniente che l'IA si sostituisca al medico, se l'IA sia affidabile, se ad essa si possa conferire una qualche sorta di responsabilità, se il paziente debba essere informato di star venendo curato con l'ausilio dell'IA, ed altri ancora.

Riguardo alla sostituzione del medico con i sistemi di intelligenza artificiale, il tema è indubitabilmente complesso. Non è da escludersi che in futuro vengano elaborate macchine che superino il test di Turing anche in ambito medico, e quindi la sostituzione tra uomo e macchina potrebbe essere possibile. Ma sarebbe davvero conveniente? Vi sono, come in ogni cosa, dei punti a favore e dei punti contrari. Certamente le macchine non rischiano di vedere alterate le proprie funzionalità a causa della stanchezza o dello stress, come invece può accadere ad un medico al termine di un turno prolungato. D'altro canto il medico, in quanto umano, possiede quella capacità di uscire dagli schemi, e quegli aspetti creativi che in molti casi si sono rivelati rilevanti per salvare la vita a dei pazienti, ma di cui, ad oggi, le macchine non sono dotate. La scelta vincente sembrerebbe dunque **l'integrazione tra le notevoli capacità delle macchine, e le ancor più notevoli capacità umane.** Anche nel merito dell'**affidabilità dell'IA**, bisogna fare delle distinzioni. Sebbene infatti si dimostri affidabile qualora operi nei limiti per cui è stata programmata, qualora invece le si richieda di operare al di fuori di quei limiti, non è più in grado di agire correttamente. Manca di intuizione e di creatività, a differenza dell'uomo. Qualora addirittura si fosse eseguito un errore durante la programmazione, la macchina agirebbe seguendo direttive sbagliate, il che, nel caso della medicina, può essere molto pericoloso. E soltanto un essere umano può rendersi conto dell'errore, e correggerlo. Nuovamente, la soluzione più semplice sarebbe un **lavoro sinergico tra IA e uomo**.

Ma qualora la macchina dovesse commettere degli errori che abbiano ricadute sulla salute dei pazienti, di chi sarebbe **la responsabilità**? È la macchina la responsabile? È forse chi l'ha programmata? È forse il medico, che potrebbe averla usata male? Bisogna fare le dovute distinzioni, in base a dove sia posto l'errore, perché c'è sia chi produce le macchine, sia chi le usa. Ovviamente, qualora vi sia un errore di programmazione, il responsabile è il programmatore. Tuttavia i programmatori sanno bene quanto sia difficile trovare l'errore nel codice di programmazione, e dunque quanto sia difficile stabilire se vi sia effettivamente un errore di programmazione, o se l'errore sia di altro genere, magari un errore delle parti meccaniche. E ancora, il venditore del software e della macchina, prima di venderlo, dovrebbe sincerarsi che sia perfettamente e correttamente funzionante. Vi sarebbe dunque anche la responsabilità del venditore. Qualora invece vi sia stato un errore di utilizzo della macchina, la responsabilità va attribuita a chi lo ha causato, che nel caso della medicina è il medico. Tuttavia nel caso di queste macchine non esistono ancora norme specifiche che regolino questo aspetto. In ultima battuta, è opportuno esaminare se il paziente debba essere informato dell'utilizzo dell'intelligenza artificiale per la sua terapia. La legge sul **consenso informato** impone al medico di spiegare al paziente tutti gli aspetti della terapia a cui il paziente sta per sottoporsi, dunque anche l'eventuale impiego dell'IA. Tuttavia per spiegare al paziente, talvolta anziano, talvolta confuso, il tipo di IA adottata nel suo caso, e come essa funzioni, può essere difficile.

Tuttavia in questo caso a venirci in aiuto è l'IA stessa, che ha dimostrato notevoli applicazioni nel campo della comunicazione, riuscendo spesso a trovare il modo più diretto e comprensibile con cui veicolare un messaggio.

In conclusione, **l'integrazione dell'intelligenza artificiale** nell'ambito medico presenta una serie di sfide etiche che richiedono un approccio oculato e riflessivo. Se da un lato l'IA può offrire soluzioni innovative e migliorare l'efficienza dei trattamenti medici, dall'altro è fondamentale affrontare con responsabilità e sensibilità le questioni legate al rapporto tra essa ed il personale sanitario, i pazienti, alla trasparenza degli algoritmi, alla responsabilità morale nelle decisioni cliniche. È essenziale promuovere una stretta collaborazione tra professionisti medici, esperti di etica, sviluppatori di tecnologia e regolatori per garantire che l'IA sia utilizzata nel rispetto dei valori fondamentali della medicina e nell'interesse del benessere umano. Solo attraverso un **dialogo aperto e un impegno condiviso** sarà possibile massimizzare i benefici dell'IA in ambito medico, mitigando al contempo i rischi etici che essa comporta.

Lo sviluppo dell'IA dovrebbe seguire dunque un **codice etico**, ossia una regolamentazione che delinei i limiti etico-morali in cui lo sviluppo dell'IA può muoversi, stilato dall'ONU, così come si riporta di seguito:

Principi Etici per l'Intelligenza Artificiale approvati dall'ONU

1. **Umanità:** L'IA dovrebbe essere sviluppata e utilizzata per migliorare il benessere umano, rispettando la dignità, i diritti, la libertà e la diversità di tutte le persone.
2. **Giustizia:** Le tecnologie IA dovrebbero essere equamente distribuite, garantendo che tutti possano accedervi e beneficiarne senza discriminazioni o ingiustizie.
3. **Trasparenza**: Le decisioni e le azioni dell'IA devono essere comprensibili e tracciabili, consentendo alle persone di comprendere le logiche e le implicazioni delle decisioni prese dall'IA.
4. **Responsabilità:** Gli sviluppatori e gli utilizzatori di IA devono assumersi la responsabilità delle azioni e delle conseguenze delle tecnologie IA, garantendo che siano conformi agli standard etici e legali.
5. **Non danneggiare:** L'IA non dovrebbe essere utilizzata per causare danni fisici o psicologici alle persone, né per violare i diritti umani o indebolire la sicurezza individuale o collettiva.
6. **Privacy e Dati**: L'IA dovrebbe rispettare e proteggere la privacy delle persone e garantire la sicurezza e l'integrità dei dati personali.
7. **Ambiente e Sviluppo Sostenibile**: Lo sviluppo e l'uso dell'IA dovrebbero contribuire alla protezione dell'ambiente e alla promozione dello sviluppo sostenibile.

Il **ruolo dell'Europa** si è rivelato fondamentale in quanto ha regolamentato, per gli stati membri, l'utilizzo e lo sviluppo dell'IA, mettendo al centro l'essere umano e i valori della civiltà europea.

Questa visione comune aumenta le opportunità, combatte le discriminazioni, e tutela i diritti dei cittadini europei. La legge approvata il 13/03/24 è la prima **legge vincolante** al mondo atta a ridurre i rischi ed aumentare le opportunità sostenendo l'innovazione tecnologica.

La scuola si pone come soggetto istituzionale fondamentale nella formazione della società del futuro in quanto: informa, promuove una corretta informazione e lo sviluppo di un pensiero critico sulle opportunità e sui rischi dell'IA, sui principi etici e sulla loro applicazione.

La scuola può essere promotrice del cambiamento, se il mondo vuole garantire che l'IA non aggravi le disuguaglianze esistenti, sarà sempre più importante che ogni cittadino abbia l'opportunità di sviluppare una solida comprensione dell'IA: cos'è, come funziona e come

potrebbe avere un impatto sulla propria vita. Questo aspetto viene talvolta definite **"alfabetizzazione all'IA".** Per questo, gli insegnanti giocheranno un ruolo fondamentale. Aiutare gli studenti a imparare a vivere efficacemente in un mondo sempre più influenzato dall'IA richiede una pedagogia che ponga maggiormente l'accento sulle competenze umane (ad esempio **pensiero critico, comunicazione, collaborazione e creatività**) e sulla capacità di collaborare con gli strumenti pervasivi dell'IA nella vita, nell'apprendimento e nel lavoro.

Gli **elementi digitali**, invece di aiutare i ragazzi a crescere e maturare la loro soggettività contribuiscono ad allontanarli dalla realtà fisica e dai veri problemi. **La tecnologia diventa un filtro deformante che lavora per isolarli dal mondo reale, dalle sue difficoltà**, delle sue incertezze, delle sue imperfezioni potrebbe generare **una immaturità perenne**, pericolosa per loro e per la società. I grandi centri commerciali, le sale d'ingresso dei grandi palazzi moderni, le grandi stazioni o gli aeroporti sono classici esempi di "non luoghi". Spazi in cui masse di individui si vengono a trovare senza interagire, senza parlarsi, spesso senza guardarsi, presi come sono dalla frenesia della tecnologia, anche un'aula scolastica si può trasformare in **"un non luogo"** se l'alunno si assenta con l'uso del cellulare. Nel non-luogo l'essere umano perde la sua personalità e il suo essere persona.

IA possiede molte potenzialità, lati oscuri e lati luminosi sta all'uomo, all'educatore far emergere le positività.

Nella società odierna e maggiormente in quella del futuro, saremo sempre più coinvolti direttamente o indirettamente con l' IA. Non scegliere una tecnologia, vuol dire subirla e quindi non poter essere inseriti pienamente nella realtà futura.

"Un giorno le macchine riusciranno a risolvere tutti i problemi,
ma mai nessuna di esse potrà porne uno."

Albert Einstein

Redazione *Segreteria del Comitato Nazionale per la Bioetica*

Comitato Nazionale per la Bioetica
Via della Mercede n.96 – 00187 - Roma