

# IA-scillante: Della Soglia Del Tempo In Cui Ci Troviamo

*Tetiana Haidash*

Ci sono delle domande che sono importanti per l'individuo, per la società e per l'umanità nel suo insieme. Dalle quali la civiltà ha preso il suo inizio.

Un esempio può esserne la scritta nel tempio di Apollo a Delfi — ΓΝΩΘΙ ΣΑΥΤΟΝ, ovvero "conosci te stesso".

Il concetto di 'conoscere se stesso' è proprio un esempio di tale questione, che ha dato inizio alla filosofia e alla scienza, e ha contribuito alla formazione della civiltà europea.

Oggi queste domande vengono affrontate in modi completamente nuovi, soprattutto a partire dai campi dell'intelligenza artificiale e dalla ricerca sulle scienze cognitive.

Abbiamo la coscienza? Abbiamo una personalità? Siamo un'illusione di noi stessi?

Questi e molti altri concetti simili sono stati discussi attivamente negli ultimi decenni, soprattutto a partire dalla cosiddetta svolta cognitiva degli anni '50.

Mai prima d'ora l'uomo si era trovato così vicino a un'entità artificiale che, pur priva di corpo, dispone di una potenza mentale tanto grande quanto la sua, se non di più. E mai prima così tanto come adesso si era messa in discussione la presenza di coscienza in un computer, mentre siamo ancora incapaci di definire e comprenderla in un essere umano.

La coscienza nella mente umana è ancora un enigma che sin dall'inizio, ha attirato l'attenzione dei più grandi pensatori dell'esistenza umana.

Ma per quanto riguarda quelli che si occupano di entrambe problematiche insieme e che vorrei trattare in questo articolo, sono soprattutto António Damásio, John Searle e Yuval Harari.

Inoltre verranno menzionati anche di Ludwig Wittgenstein, Gilbert Ryle, Michio Kaku e Jean Baudrillard.

Vorrei iniziare da António Damásio, neurologo, neuroscienziato, psicologo e saggista portoghese. Nel suo libro "Il sé viene alla mente. La costruzione del cervello cosciente" Damasio vede la coscienza come un prodotto dell'evoluzione del nostro cervello ed esplora quali processi nel cervello portano all'emergere della coscienza.

Discute anche la teoria del libero arbitrio, sostenendo che le nostre decisioni e azioni possono essere determinate dal nostro cervello, dall'ambiente e da altri fattori, proponendo la nozione di libero arbitrio illusoria.

Ma la sua idea riguardante IA la esprime nel suo libro "L'ordine delle cose", nella quale avverte che l'intelligenza artificiale non può sostituire completamente l'intelligenza umana; le macchine non hanno le stesse sensazioni emotive e fisiche che hanno gli esseri umani.

Parla anche della sua preoccupazione riguardante la possibilità di utilizzare l'intelligenza artificiale per il controllo delle persone e limitazione della loro libertà. Cosa, che diventa sempre più soggetto ai dubbi e discussioni, tanto da mettere al dubbio la sua utilità.

In un articolo su *JAIR (Journal of Artificial Intelligence Research)* dove viene trattato proprio questo tema viene detto che uno dei modi per prevenire le conseguenze dannose delle azioni di una intelligenza artificiale è il suo completo contenimento o controllo su quali azioni può compiere.

Però è impossibile ottenere un controllo completo con i metodi moderni. E questo è logico, perché per controllare completamente un intelletto di tale potenza, è necessario un intelletto altrettanto potente. Il che porta alla natura ciclica di questo problema.

Ma nonostante questo, Damásio certamente sottolinea il fatto che l'intelligenza artificiale può essere uno

strumento utile in molti settori, tra cui la medicina e la scienza. A questo proposito c'è stato fatto un interessante esperimento mentale da John Searle chiamato *Stanza Cinese* e pubblicato nel 1980 sulla rivista *"The Behavioral and Brain Sciences"* nell'articolo *"Minds, Brains and Programs"*.

Si pensa che un computer può sostituire una mente, in altri termini può diventare esso stesso una mente una volta programmato e quindi pensare. Al contrario, John Searle vuole dimostrare che una macchina può soltanto simulare una mente in quanto non ha condizioni per poter pensare. Quindi i programmi dei computer sono puramente sintattici e non semantici; possono produrre collegamenti corretti, costruire frasi corrette o operare correttamente con simboli, ma non possono dare significato a questi simboli, non sono collegati dal senso e di conseguenza i computer non hanno una vera coscienza.

Ma adesso potrebbe sorgere una possibile controargomentazione, che potrebbe essere formulata ponendosi la domanda se, ad esempio, uno studente riesca a comprendere appieno il significato della propria risposta quando il professore gli chiede qualcosa?

Ad esempio che cosa sia l'apercezione trascendentale di Immanuel Kant? O quali sono le idee di Platone?

In linea di principio, questo sta già accadendo nei social dove persone che si considerano delle personalità e si scambiano con una serie di *cliché*; ma non vengono accusati ad esempio per non rendersi conto di ciò che dicono.

Cioè, tali richieste non vengono fatte a questi individui. Non viene chiesto loro se sono consapevoli di ciò che stanno affermando, come fanno a sapere cosa affermano, se possono giustificare quello che dicono?

In questo, il programma sa già farlo meglio delle persone sui social.

A questo punto inizia a sorgere la domanda se, in qualità di esseri umani, sia necessario comprendere e percepire ciò che esprimiamo?

Tutti gli esperimenti filosofici ed intellettuali hanno al centro proprio questo interrogativo, espresso attraverso miti come quello della caverna di Platone, della stanza cinese di Searle, della stanza di *Mary Jackson* o del mulino di *Leibniz*...

Tali finzioni intellettuali mirano a chiarire se le nostre affermazioni siano frutto di esperienza o illusione. Se siamo realmente consapevoli di ciò che diciamo e affermiamo.

Michio Kaku, rinomato fisico giapponese e divulgatore scientifico, come risposta all'interrogativo di Searle si chiede se una macchina debba realmente comprendere il significato dei caratteri (in questo caso quelli cinesi come nell'esperimento), sottolineando invece l'importanza che essa risponda in maniera corretta.

Curiosa è però anche la concezione di Gilbert Ryle, uno dei maggiori e più influenti esponenti della filosofia analitica, riguardo l'errore categoriale, che tra l'altro secondo lui è quello che ha originato il dualismo cartesiano tra mente e corpo.

Si tratta di un errore consistente nel "presentare i fatti della vita mentale come appartenenti a un tipo o categoria (o classe di tipi o categorie) logico (o semantico) diverso da quello cui essi invece appartengono".

Succede ad esempio quando viene detto che l'università ha la conoscenza.

È evidente che si intende l'università come luogo in cui la conoscenza viene generata, immagazzinata e trasferita e non invece l'edificio stesso che le ha.

Analogamente possiamo dire che i computer possono pensare. Ma in realtà, il pensiero è una caratteristica che è unica per gli esseri umani.

I computer possono eseguire determinati compiti ed elaborare informazioni, ma non sono in grado di pensare, poiché il pensiero è un processo associato alla comprensione, alla percezione, alla valutazione e al prendere decisioni basate sull'esperienza e sulla conoscenza personali.

Curiosamente, la ciclicità si ripete e viviamo in un'epoca in cui vecchie questioni filosofiche diventano sempre più attuali, come ad esempio: cosa significa conoscere? Cosa significa conoscere se stessi? Cosa significa

comprendere se stessi, gli altri e il mondo? Cosa significa pensare, avere coscienza, essere una personalità, avere la volontà libera e prendere decisioni consapevoli?

Il problema è che la filosofia nel XX secolo, intenzionalmente, ha fatto calare l'importanza di queste questioni e altre simili.

Per Wittgenstein, il filosofo austriaco-britannico che ha studiato il linguaggio, la lingua non descrive nulla al di fuori del linguaggio stesso. Il significato delle parole è legato ai contesti in cui vengono utilizzate, ai giusti usi delle parole (giochi del linguaggio).

Non è così come pensavano ad esempio John Locke, Tommaso d'Aquino o Platone.

Tutto ciò che accade nello spazio quindi è il risultato di scambi linguistici.

Secondo Wittgenstein, le nostre pratiche (il linguaggio è parte delle nostre pratiche) possono essere descritte come un insieme di partite a scacchi. E poiché le partite a scacchi possono alla fine essere ridotte a un punteggio e diventare un potente programma - come dimostra la storia degli ultimi decenni - allora possiamo anche ricreare il linguaggio, la coscienza, il pensiero e la comprensione sotto forma di programmi.

In una delle opere di Gilbert Ryle, egli cerca di dimostrare che in realtà esiste una priorità della conoscenza pratica, ovvero quella che chiamiamo il mondo interiore, l'esperienza del nostro io, i nostri vissuti interiori, quindi il processo di interiorizzazione.

Jean Baudrillard, sociologo e filosofo francese noto per la sua critica della società moderna, della cultura del consumo e della cultura di massa in generale.

Baudrillard è conosciuto anche per le sue idee sui simulacri, dove sostiene che il mondo moderno non è pieno di realtà, ma simulacri, cioè copie di copie che non sono più correlate agli oggetti originali.

Lui non aveva alcun legame diretto con la robotica, ma le sue idee furono importanti per lo sviluppo della teoria dell'intelligenza artificiale e della cibernetica. Nei suoi scritti considerava i problemi associati al progresso tecnologico e sosteneva che la tecnologia, compresa quindi l'intelligenza artificiale e i robot, può aumentare il dissenso tra realtà e simulacri.

Ma dal momento che Baudrillard ha espresso i suoi pensieri riguardo queste tematiche sono passati molti anni, e in relazione al ritmo che si è sviluppato negli ultimi decenni e che si sta solamente velocizzando, è difficile prenderlo in considerazione.

Ogni giorno abbiamo accesso a una moltitudine di informazioni, ma nel tentativo di comprendere ciò che sta accadendo, acceleriamo l'accumulo di conoscenze, provocando così cambiamenti sempre più rapidi e su scala sempre maggiore. Alla fine, diventiamo sempre meno capaci di comprendere il presente e di prevedere il futuro. Più sappiamo, meno possiamo prevedere.

Durante la nostra storia, la tendenza a inventare qualcosa per uno scopo specifico che con il tempo diventava ordinarietà — come molte altre cose — si è ripetuta più volte.

Questo è particolarmente evidente nel campo della medicina, dove non ci sono confini definiti. Quando si raggiunge un risultato rivoluzionario nella cura, non si può vietare di utilizzare questa metodologia per una trasformazione nuova.

Ad esempio, quando il *Viagra* è stato introdotto per la prima volta come farmaco per normalizzare la pressione sanguigna, è stato rapidamente scoperto che aiutava gli uomini con problemi di impotenza, il che è stato un vantaggio. Tuttavia, presto lo hanno iniziato ad usare anche gli uomini senza questi problemi, solamente per la curiosità di capire quanto fosse possibile superare la propria norma.

Così è accaduto anche con la chirurgia plastica, che inizialmente aveva solamente lo scopo di massimizzare il ripristino di parti del corpo umano durante la Prima Guerra Mondiale. Quello che è successo dopo lo sappiamo già.

Queste e molte altre prove simili dimostrano quanto sia impossibile arrivare alla soddisfazione unica e la necessità di sviluppare utilizzi sempre nuovi. Per cui noi capiamo che costruendo ad esempio gambe bioniche che permettono a persone con paralisi degli arti inferiori di camminare di nuovo, questa tecnologia sarà

utilizzata anche per un *upgrade* di persone sane.

Quindi è probabile che il processo di trasformazione sarà un processo storico e graduale.

*Homo sapiens* non sarà sterminato dai robot ribelli. Piuttosto, lentamente si modificherà, fondendosi sempre di più con i robot e i computer, fino a quando i nostri discendenti non si guarderanno intorno e si accorgeranno di non essere più la stessa creatura.

Questo non avverrà in un giorno e neanche in un anno. In realtà sta già accadendo ora, attraverso molte azioni comuni. Ogni giorno, milioni di persone concedono ai loro smartphone un'altra parte della loro vita o prendono un antidepressivo più efficiente appena creato.

Nella ricerca della salute, della felicità e del potere, l'uomo cambierà prima una sua caratteristica, poi un'altra, una terza; fino a smettere di essere un umano.

Per millenni la storia è stata scossa da cambiamenti tecnologici, economici, sociali e politici, ma una cosa è rimasta immutata: l'umanità stessa.

Ora stiamo imparando a ricostruire i corpi e i cervelli principalmente per evitare l'invecchiamento, la morte e la "sofferenza", ma una volta acquisita questa abilità, chissà su cosa decideremo di usarla ancora?

Siamo consapevoli che circondati da delle super persone, la nostra personalità, i nostri sogni e persino le nostre paure diventeranno insignificanti e non ci sarà più alcun beneficio per noi. Questo è un timore che collettivamente condividiamo come specie umana adesso ed è normale.

Colui che realmente crede che saremo in grado di fermarci — anche se ha una comprensione accurata del funzionamento del cervello e dei computer — probabilmente ha una scarsa conoscenza della psiche umana e della società umana.

Secondo Harari, la mente è un flusso di esperienze soggettive che consistono in sensazioni, emozioni e pensieri che, intrecciandosi, diventano il flusso della coscienza.

I robot non hanno sensazioni e non desiderano nulla, di conseguenza non hanno coscienza. Attualmente, la scienza sa sorprendentemente poco su come funziona la mente umana. Finora, nessuno ha capito come le correnti elettriche nel cervello creano sensazioni soggettive e perché l'uomo ha bisogno di sperimentare la fame o la paura.

È sorprendente che più gli scienziati comprendono come funziona il cervello, più misteriosa diventa la funzione della mente e dei sensi.

Ancora non si sa se accade qualcosa nella nostra mente che sia diverso da ciò che accade nel cervello. Tuttavia, nessun sistema di elaborazione dell'informazione ha bisogno di sensazioni soggettive per funzionare normalmente.

Se non possiamo spiegare come funziona la nostra coscienza, allora dovremmo semplicemente ignorarla?

Per millenni, le persone hanno spiegato i fenomeni naturali incomprensibili come la volontà divina. Tuttavia, adesso nessuno più di tanto prende sul serio l'esistenza di Dio. Lo stesso destino è toccato al concetto di anima...

Gli scienziati concordano sul fatto che la coscienza esiste e ha valore morale, ma non svolge alcuna funzione biologica. Forse la coscienza è un sottoprodotto biologicamente inutile dei processi che avvengono nel cervello.

Fino a poco tempo fa, si credeva che solo gli esseri coscienti potessero svolgere compiti altamente intelligenti. Ma lo sviluppo della cibernetica e delle tecnologie informatiche indica il contrario: l'esecuzione di compiti complessi richiede certamente alta intelligenza, ma la presenza di coscienza non è necessaria.

Come giustamente osservato da Harari nel suo libro *Homo Deus: Breve Storia del Futuro* "la scelta di questi progetti potrebbe essere un errore grandioso. Ma la storia è piena di errori grandiosi", aggiungendo che "a giudicare dai nostri successi passati e dai nostri valori attuali, continueremo comunque a cercare la felicità, la divinità e l'immortalità — anche se alla fine ci ucciderà".

Una parte significativa dei pensatori contemporanei afferma che la coscienza sia un'illusione, ma allo stesso tempo cerca di attribuire la coscienza ad un programma.

Affermano che la personalità umana sia un'illusione, ma nel contempo cercano di parlare della personalità della macchina.

Parlano del nostro libero arbitrio e delle nostre decisioni come di illusioni, ma sostengono che gli animali hanno libertà di volontà, consapevolezza, giustizia, altruismo e molte altre virtù etiche e morali.

Ecco in che modo si sviluppano le discussioni oggi, in questi contrasti.

Tuttavia, tutte queste questioni sono estremamente importanti perché ci portano a domande fondamentali della filosofia e della scienza, ci costringono a riflettere sulla natura della coscienza, della personalità, sul significato di avere una biografia, di svilupparsi, di porsi obiettivi di vita e di strategie di sviluppo dell'umanità.

## **Riferimenti bibliografici:**

TED. *"The Quest to Understand Consciousness"* YouTube, uploaded by TED, 19 Dec. 2011, [https://youtu.be/LMrzdk\\_YnYY](https://youtu.be/LMrzdk_YnYY)

TED. *"Consciousness and The Brain"* YouTube, uploaded by TED, 24 May. 2013, [https://youtu.be/j\\_OPQgPIIdKg](https://youtu.be/j_OPQgPIIdKg)

Andrii Baumeister. *"Una Macchina Ha un'Anima? Come Può l'Intelligenza Artificiale Risvegliarsi Alla Vita?"* YouTube, uploaded by Andrii Baumeister, 21 Jan. 2023, [https://youtu.be/iobc\\_I3nmj4](https://youtu.be/iobc_I3nmj4)

Harari, Y. N. (2018). *Homo Deus: A brief History of Tomorrow*. Sindbad

Damasio, A. (2012). *Il Se Viene Alla Mente: La Costruzione del Cervello Cosciente*. Random House

Alfonseca, Manuel. *"Superintelligence Cannot be Contained: Lessons from Computability Theory"*. JAIR, published by JAIR, 5 Jan. 2021, <https://www.jair.org/index.php/jair/article/view/12202>

Maksutov, Ivar. *"Intelligenza Artificiale Sostituirà l'Uomo?"*. Post Nauka, published by Postnauka, 27 July. 2021 <https://postnauka.ru/tv/156532>

Maksutov, Ivar. *"Algoritmi ed Etica: l'Intelligenza Artificiale Migliorerà la Nostra Società?"*. Post Nauka, published by Postnauka, 26 August. 2020 <https://postnauka.ru/tv/155689>