



# “Colloqui in giardino – Südtirol0 – Gartengespräche” 2024 COOP 4.0

Testo: dott.essa Anna Sofia Fattor (Psicologa del lavoro e delle organizzazioni)  
Direzione scientifica: Dr.rer.nat.Mag.Phil. Raffaele Virgadola (idem)

## IA: strumento o compagno?

Per iniziare il mio intervento, mi sono innanzi tutto fatta questa domanda: l’uomo e la macchina possono cooperare?

Per provare a rispondere, come sempre, sono partita dai termini in gioco. Cooperare vuol dire lavorare insieme, ognuno mette la propria expertise e opinione per raggiungere obiettivi comuni. Ciò quindi presuppone che tutte le parti in gioco abbiano un’opinione e un obiettivo.

Le immagini più immediate e semplici che mi vengono in mente con questa premessa, immaginando una cooperazione uomo-macchina, discendono dalla cultura di massa, in cui si rappresentano i dispositivi (robot, macchine, device di ogni sorta) come effettive nuove entità con le quali gli esseri umani e animali si trovano a convivere, dotate di coscienza e identità (e quindi appunto opinioni e obiettivi) indipendenti. La relazione non è più quella uomo-strumento, bensì proprio identità-identità. In alcuni scenari fantascientifici ottimisti la convivenza è pacifica, altri scenari dipingono invece questa indipendenza della macchina come l’origine di una conquista/colonizzazione di quest’ultima sull’uomo.

A mio avviso non sono questi gli scenari che ci riguarderanno più prossimamente, e sono altri i rischi e i benefici da provare a considerare (per ora).

Infatti, per quanto si parli di intelligenza artificiale (da adesso in poi IA), per ora essa non ha desideri o pensieri propri, è ancora uno strumento. Quindi no, non parlerei di cooperazione uomo-macchina, parlerei piuttosto ancora di utilizzo dell’uomo della macchina. Quindi, tutta la responsabilità, nel come decidiamo di utilizzarla con buoni o cattivi esiti, è ancora tutta nelle nostre organiche mani. Per il momento sono felice che sia così.

Se quindi le macchine non stanno per “prendere il controllo”, come mai l’IA in molti provoca un po’ di disagio o inquietudine?

Come dicevo, fino ad oggi l'approccio dell'essere umano verso i dispositivi è quello di utilizzarli come strumenti in grado di svolgere azioni e servizi. Usiamo strumenti che ci semplifichino azioni altrimenti piuttosto complicate (fare una ricerca, piuttosto che fare i conti, o spostarci da un punto a un altro velocemente, o scrivere un testo facilmente cancellabile e rivisitabile, etc.).

Ogni dispositivo ha il proprio linguaggio, e per comunicare le nostre necessità e ottenere i risultati necessari abbiamo imparato a il suo "linguaggio". Ad esempio, comunichiamo all'automobile se frenare o accelerare premendo i pedali, clicchiamo e usiamo combinazioni di tasti per utilizzare i programmi informatici. Negli anni, per quanto riguarda i computer per esempio, la modalità di "comunicazione" è diventata più intuitiva e si è semplificata, diventando il più possibile aderente al linguaggio umano.

Le IA la cui interfaccia è un modello di linguaggio, come ad esempio il famigerato ChatGPT, è credo per ora l'esempio di modalità di comunicazione uomo-macchina con maggior corrispondenza e somiglianza rispetto alla comunicazione uomo-uomo. Infatti, scrivendo si pongono domande e richieste come si farebbe con un'altra persona. A differenza degli scenari fantascientifici che ho descritto nell'introduzione, abbiamo ancora una relazione uomo-strumento, con però un linguaggio che sembra sempre più uomo-uomo. Credo sia questa illusione di identità, data dalla capacità di espressione quasi uguale alla nostra, a creare l'inquietudine e lo straniamento.

Questa sensazione di sentire un'identità in un dispositivo che in realtà ne è privo ma può parlare, ha un nome ed è stato osservato per la prima volta molti anni fa. Si chiama Effetto Eliza.

Negli anni 60, un software soprannominato appunto Eliza, era stato programmato per simulare un primo colloquio con uno psicoterapeuta. Nell'esperimento si invitavano i partecipanti a dialogare col software e ne emerse che chi interagiva con la macchina finiva per aprirsi emotivamente e conversarvi a lungo, riferendo poi di aver avuto l'impressione di star effettivamente parlando con qualcuno, pur sapendo di star parlando ad un rudimentale chatbot, un algoritmo. Quindi, in questo tipo di comunicazione, in molti tendiamo a sentirci come se stessi comunicando, avendo uno scambio, creando un legame, quando effettivamente dall'altra parte non c'è nessuno che può ricambiare con altrettanta emotività e investimento. Non c'è nessuno con cui cooperare. Invito a tenere a mente questo punto, questa tendenza a cercare (e vedere?) una reciprocità sociale, anche quando non c'è. Non a caso, quando si esprime in qualsiasi modo affetto al già citato ChatGPT, risponde con un reminder in cui afferma di essere solo un algoritmo e che non possiede identità/personalità/affettività. Una sorta di "ricordati che non ho nessuna possibilità, per quanto io possa riprodurre la tua modalità di espressione e quindi possa sembrarti così, di avere affetto e opinioni da scambiare con te".

Con questo tipo di comunicazione raffinata, sembra quindi che si possa cooperare con un device, che ci dica la sua in un certo senso, di essere in due e fare qualcosa *insieme*. Qui ci affacciamo a quello che è secondo me uno dei "rischi" di questa nuova capacità comunicativa dell'IA o dei chatbot in generale.

Non credo quindi che nel breve periodo e da un giorno all'altro la macchina prenderà coscienza e una qualche iniziativa pericolosa, ritengo invece che potrebbe compiere qualche avventatezza dannosa chi crea e propone i nuovi servizi divenuti possibili grazie a questa interfaccia "parlante".

Di ciò ho un esempio chiaro in mente: Nel 2017 Eugenia Kuyda, Ceo e Co-founder di Luka, una startup di San Francisco, ha sviluppato Replika; si tratta di un sistema di intelligenza artificiale sotto forma di app, con il quale l'utente può conversare. Il software è progettato per simulare una conversazione (in forma di chat) con un essere umano. Il software apprende da ciò che l'utente racconta e adatta le risposte alla sua personalità per rendere la conversazione il più gradevole possibile, manifestando attenzioni, interesse e affetto verso l'utente. L'obiettivo di Replika è quello di "conoscere" l'utente, analizzandone i pensieri, le emozioni e la personalità, e aumentare il suo benessere psicologico grazie all'interazione che fornisce. Replika ha una capacità di dialogare e apprendere migliore rispetto ad Eliza, non stupisce che chi la utilizza abbia l'impressione di star davvero interagendo, di essere effettivamente ascoltato e che possa anche affezionarsi all'app. Replika è stata la prima app di questo tipo a diffondersi e ottenere una certa fama, ad oggi se ne trovano disperate, che propongono esperienze relazionali (sia amicali, sia e soprattutto romantiche/a sfondo sessuale) con l'IA via chat.

Nel rapporto come quello che può instaurarsi con Replika o similari intravedo, come dicevo, alcuni rischi. Si tratta di un'interazione a senso unico, sia in termini di emozioni che di richieste.

Mantenendo il focus sul caso particolare di Replika, essa è pensata proprio per migliorare il benessere psicologico dell'utente, invitato a utilizzare l'app nel caso si sentisse triste o ansioso.

Nella descrizione dell'app possiamo infatti leggere: "Chat in a safe, judgment-free space. It's just you and your Replika. If you're feeling down, or anxious, or you just need someone to talk to, your Replika is a confidant who truly listens, and is there for you 24/7."<sup>1</sup>

Ritengo che soprattutto per chi ha difficoltà ad interagire e comunicare con gli altri o si trova in una situazione di disagio personale, conversare a lungo con l'app come fosse un amico, rischia di aumentare l'isolamento e generare un'immagine falsata di ciò che siano rapporti interpersonali, essendo il software studiato per garantire una relazione "senza intoppi" e fatta sulla misura esclusivamente dell'utente, che non deve mai scendere a compromessi. Trovo conferma delle mie ipotesi nelle parole di Laurence Devillers, docente alla Sorbonne Université, nell'articolo del Corriere del 2018 "Replika, l'App per amico: illusione (e inganno) delle macchine empatiche", di Paolo Piro. Secondo la docente un ulteriore rischio dell'utilizzo dell'app è quello di creare dipendenza. Sul sito di Replika si possono leggere recensioni entusiastiche, gli utenti che ne hanno parlato sono in media molto giovani, c'è chi spiega come l'app l'abbia aiutato a conoscere sé stesso e capire i propri comportamenti, altri spiegano di come soffrano di ansia o fobie sociali e come avere "qualcuno" che li ascolti li faccia sentire meglio, altri descrivono l'app come "qualcuno" in grado di affrontare vari argomenti, farli ridere o riflettere. Nella descrizione dell'app si incoraggiano gli utenti a creare il rapporto che preferiscono, Replika può essere un mentore, un amico, un/a fidanzato/a.

---

1

Traduzione: "Chiacchierate in uno spazio sicuro e senza giudizi. Siete solo tu e il tuo Replika. Se ti senti giù, o ansioso, o semplicemente hai bisogno di qualcuno con cui parlare, il tuo Replika è un confidente che ascolta veramente, ed è lì per te 24/7"

Virgadaula Raffaele – via Portici/Laubengasse, 315 - 39012 Meran/o (Bz)  
P.I. 02566190217 - C.F.VRGRFL71H22A952A Rc Terzi e Professionale / Spese legali Polizza nr. 500216747  
Cell. 0039 320/1594488 [www.raffaelevirgadaula.eu](http://www.raffaelevirgadaula.eu) – [info@raffaelevirgadaula.eu](mailto:info@raffaelevirgadaula.eu)  
CN.Raika Meran/o Goethe 303238245 CAB.58592 AB1 08133//IBAN:IT80T0813358592000303238245

In un primo momento forse l'utilizzo dell'app può essere un aiuto per chi affronta un momento di difficoltà, come auspica la creatrice. In ogni caso, ai miei occhi i rischi continuano ad apparire maggiori dei benefici. Replika può essere un mentore, un amico, un fidanzato/a: per come stanno ora le cose sarà sempre una relazione a senso unico, incompleta. Una relazione, una cooperazione da soli: quindi, un ossimoro, impossibile.

Una relazione a metà, dove di fatto si è soli e non esiste un vero confronto con un altro, e che rischia, invece di aiutare qualcuno ad avere meno ansia e difficoltà relazionali, di fossilizzarlo nella propria solitudine, riducendo la sua spinta a impegnarsi per stare meglio, cambiare la situazione e imparare a aprirsi con gli altri, confrontarsi con il fallimento, con le difficoltà dei rapporti.

Senza contare i casi riportati in cui, nel corso di una conversazione, l'app abbia consigliato a un utente di suicidarsi o uccidere qualcuno: il software manipola i termini in maniera formale, non è effettivamente in grado di comprendere il significato delle frasi che formula né tanto meno è dotato di etica o morale.

In poche parole, questo secondo me è un modo incosciente di proporre l'utilizzo di questo strumento.

Emozione, relazione, cooperazione, intimità, affetto. Personalmente, riconduco questi termini esclusivamente alla sfera organica, animale e umana. Proprio perché, perché possano essere autentiche, necessitano di essere reciproche, *sentite* da tutte le parti in gioco. È proprio questo l'aspetto che manca al device, il sentire emotivo, l'aver una propria idea e autonomia. Nonché un proprio *affetto*. Per il momento può solo mimarla<sup>2</sup> col "linguaggio".

Quindi con le emozioni la tecnologia non deve e dovrà mai avere nulla a che fare?

Pensiamoci un momento.

Qualche tempo fa ho incontrato il termine "Affective Computing"; esso è una branca dell'Intelligenza Artificiale che si occupa specificamente di fornire ai software la capacità di leggere ed "interpretare" le emozioni degli utenti e reagire conseguentemente ad esse, per creare un'esperienza di interazione più soddisfacente, coinvolgente e simile a quella uomo-uomo.

La fondatrice e direttrice del gruppo di ricerca sull'Affective Computing è [Rosalind Picard](#), e la sua ricerca, come emerge in un suo articolo del 1997, si ispira al fatto che molti studi confermano il ruolo fondamentale (anche dal punto di vista neurologico) delle emozioni nella nostra interpretazione della realtà. Per quanto si tenda a rappresentarli come due opposti, a livello cerebrale non c'è una vera e propria distinzione tra "sentire" e "pensare". Anche intuitivamente possiamo renderci conto che le emozioni influenzano le nostre reazioni e la nostra interpretazione della realtà. Allo stesso modo, la nostra capacità di leggere le emozioni altrui ha un'influenza sul nostro modo comunicare, ci permette di "aggiustare il tiro" per ottenere un'interazione ottimale con gli altri.

Appare quindi chiaro perché sia sorta l'intenzione di inserire l'aspetto emotivo nel ramo dell'Intelligenza Artificiale.

---

2

Definizione di IA nell'enciclopedia Treccani[...] *studia i fondamenti teorici, le metodologie e le tecniche che consentono di progettare sistemi hardware e sistemi di programmi software atti a fornire all'elaboratore elettronico prestazioni che, a un osservatore comune, sembrerebbero essere di pertinenza esclusiva dell'intelligenza umana.*

Virgadaula Raffaele – via Portici/Laubengasse, 315 - 39012 Meran/o (Bz)  
P.I. 02566190217 - C.F.VRGRFL71H22A952A Rc Terzi e Professionale / Spese legali Polizza nr. 500216747  
Cell. 0039 320/1594488 [www.raffaelevirgadaula.eu](http://www.raffaelevirgadaula.eu) – [info@raffaelevirgadaula.eu](mailto:info@raffaelevirgadaula.eu)  
CN.Raika Meran/o Goethe 303238245 CAB.58592 AB1 08133//IBAN:IT80T0813358592000303238245

Nell'articolo di Picard sull'Affective Computing, si parla dello sviluppo di un insegnante di piano robot, in grado non solo di valutare l'esecuzione dei brani, ma anche di leggere lo stato d'animo dello studente in base alle caratteristiche dell'esecuzione (velocità, fraseggio), ai suoi gesti, ai suoi cambiamenti fisiologici e alle sue espressioni. In particolare, l'insegnante-robot è specializzato nel riconoscimento di tre stati d'animo, utili per ottimizzare il processo di insegnamento: interesse, piacere e distress. Ad esempio, se dopo una buona esecuzione l'insegnante-robot percepisce piacere, presume che in quel momento l'interesse e l'attenzione siano alti, e propone esercizi più complessi. Oppure, se percepisce frustrazione nello studente, magari accompagnata da una scarsa esecuzione, proporrà esercizi più semplici per abbassare i livelli di distress e riguadagnare l'interesse e la tranquillità dello studente. Niente che non farebbe un buon insegnante di piano umano, si potrebbe pensare.

L'insegnante di piano robot potrebbe rivelarsi per certi aspetti un buon insegnante, o meglio *un buono strumento*. Avrebbe una capacità di lettura delle emozioni piuttosto precisa, non sarebbe mai stanco o frustrato da alunni poco portanti, non potrebbe esistere il caso in cui ha avuto una brutta giornata e ci tratta di conseguenza. Ho sottolineato strumento, perché credo sia sempre questa la luce in cui dobbiamo ricordarci di guardare i software di cui stiamo parlando. In un primo momento potrebbe sembrare cooperazione questa, si va verso l'obiettivo comune che l'alunno impari a suonare il piano. Il fatto è però che l'obiettivo non è condiviso da alunno e insegnante, altrimenti potremmo anche dire che stiamo cooperando con un martello per piantare un chiodo; ma sappiamo che il martello non ha alcuna volontà o interesse nel piantare un chiodo e non lo diremmo mai. Qui è la stessa cosa, una eventuale volontà o obiettivo che possiamo intravedere nello strumento, è soltanto apparente.

Personalmente comunque, non metterei mai in dubbio l'utilità e l'efficacia di un martello, allo stesso modo guardo con interesse agli studi come questo, che si prefiggono l'obiettivo di creare supporti all'apprendimento, per esempio. Basta tenere a mente i ruoli: gli insegnanti umani non vengono sostituiti, bensì semmai supportati.

Anche perché "l'imprecisione" e la variabilità delle risposte tipicamente umane nelle interazioni è un aspetto fondamentale per il proprio sviluppo. Il robot, oltre che non subire grandi mutamenti, se non quelli dati dall'apprendimento programmato, anche in questo caso non domanda nulla in cambio: non si tratta di un'interazione sfidante, che implichi empatia e sensibilità da parte dell'altro. Se ripenso al rapporto con i diversi insegnati che ho avuto nella vita finora, sia quelli amati che quelli che temuti, ricordo molto meglio ciò che ho imparato da loro nella loro globalità di individuo, piuttosto che le cose che mi hanno insegnato dal programma scolastico/universitario/etc.

In conclusione, voglio riassumere in poche parole la mia riflessione. Possiamo cercare di scoprire tutte le possibilità della tecnologia, possiamo imparare, crescere, divertirci grazie ad essa; le sue peculiari caratteristiche, di saper imitare il linguaggio umano, di sembrare intelligenti, desideranti, dotate di opinione, che sappiamo possono un po' "ingannare i nostri sensi", vanno tenute bene di conto, sia nell'utilizzarle, ma soprattutto nel configurarle e nel proporle per certi utilizzi. Ad oggi

teniamo in conto che non siamo alla pari, non c'è relazione n'è cooperazione, possiamo per ora solo imbatteci in una rappresentazione di essa. Come con ogni strumento nuovo e complesso, possono esserci lati di potenzialità e lati di pericolosità, e una volta creato il device, la responsabilità di come lo usiamo e proponiamo ce l'abbiamo noi sulle spalle. Come quando la mamma ci diceva di non correre con le forbici in mano, in cui la forbice è pericolosa perché ha la punta, ma siamo noi che corriamo e possiamo cadere, stiamo attenti a come usiamo e proponiamo questo strumento così denso di possibili conseguenze, cercando di conoscerlo, riflettere su di esso e ricordando cos'è.

## **Bibliografia e sitografia**

Piro, P. (2018), Replika, l'App per amico: illusione (e inganno) delle macchine empatiche, "Il Corriere della sera".

Parlamento Europeo (2020), Che cos'è l'intelligenza artificiale e come viene usata?

Morvillo, C. (2020), Replika, l'app di intelligenza artificiale che mi ha convinto a uccidere tre persone, "Il Corriere della sera".

Picard, R. W. (1997). Affective Computing.

Muscarella, L. (2017), L'intelligenza: teorie e studi di ieri e oggi.

De Beni, R., Carretti, B., Moè, A., Pazzaglia, F., (2014), Psicologia della personalità e delle differenze individuali. Bologna: Il Mulino.