

INTELLIGENZA ARTIFICIALE

Una rivoluzione digitale infarcita di Hype tra il già-visto e il totalmente-imprevedibile



Andrea Granelli,
Fondatore di Kanso

L'Intelligenza Artificiale (IA nel resto dell'articolo), nelle sue varie forme oggi in parte note, è certamente una rivoluzione resa possibile dal digitale – anche se il suo fattore trasformativo va molto oltre le funzionalità tecnologiche – che muterà in modo incisivo (e in alcuni casi anche drammatico come stiamo vedendo ad esempio nei due conflitti in Ucraina e in Palestina) il nostro modo di vivere, operare e lavorare.

L'IA non indica una specifica tecnologia, quanto un gruppo di soluzioni la cui finalità è generare comportamenti autenticamente "umani". Questa definizione, pur nella sua voluta genericità, indica in modo sufficientemente chiaro il campo d'azione di questo dominio tecnologico.

Un breve ma necessario inquadramento del problema

Non conviene addentrarsi nelle definizioni esatte perché non c'è accordo fra gli esperti e la tecnologia è ancora immatura; si cadrebbe pertanto facilmente in considerazioni filosofiche (e talvolta addirittura ideologiche) che esulano dalla finalità di questo articolo, a partire dalla definizione stessa di umano o dal fatto se sia desiderabile o meno imitare comportamenti umani e quali comportamenti emblematici vadano privilegiati.

D'altra parte, **l'imitazione della natura è sempre stata uno dei motori dell'innovazione.** Pensiamo ad esempio alla recente *bio-mimicry*, che studia e imita i processi biologici e biomeccanici della natura e degli esseri viventi come fonte di ispirazione per il miglioramento delle attività e tecnologie umane. Lo stesso digitale

è nato per imitare – migliorandoli – alcuni specifici comportamenti umani. Ha iniziato con il calcolare, il comunicare, l'archiviare, lo scrivere, il monitorare, il produrre e oggi sta entrando in modo determinato nelle aree del **consigliare**, del **valutare** e del **decidere**.

Sembrerebbe un processo in continuità con il passato, ma in verità non lo è. Automatizzare le attività da esperto – valutare, decidere e consigliare – aumenta in modo relevantissimo ed esponenziale la rischiosità di questi sistemi ed è questo il punto su cui poco si riflette.

Che la tecnologia sia un *pharmakos* – parola greca potente che indica sia il farmaco che il veleno – è un fatto assodato, anche se tenuto in scarsa considerazione. Anzi il filosofo **Bernard Stiegler** estende questa proprietà alla stessa conoscenza umana, vista come una particolare tecnologia che consente all'uomo di sopperire alle sue carenze biologiche (mancanza di pelliccia per difendersi dal freddo, sistema respiratorio non anfibio, forza corporea particolarmente limitata...) e sopravvivere in un ambiente ostile. Osserva infatti il filosofo: *"Il sapere è una cura: tutte le forme della conoscenza sono prescrizioni terapeutiche per aumentare la funzione curativa del pharmakon e ridurre la tossicità"*.

Qual è allora il punto? Come nel caso della medicina, quando prendiamo un farmaco dobbiamo avere molto chiari non solo i benefici ma anche i rischi e gli effetti collaterali. Per questo motivo i medici studiano a fondo farmacologia e patologia, cosa che invece il tradizionale addestramento digitale (l'alfabetizzazione digitale) non prevede. **I lati oscuri del digitale sono spesso totalmente**

assenti dalla formazione al digitale.

Anzi, si tende a ritenere che le nuove tecnologie digitali non siano intrinsecamente pericolose, ma lo diventino solo a causa dell'incompetenza degli utenti. Basta quindi aumentare la formazione tecnica per contenere le aree di rischio. Niente di più errato. Il rischio è strutturale e nativo. Lo dice con efficacia **Paul Virilio**: «*La tecnologia crea innovazione ma – contemporaneamente – anche rischi e catastrofi: inventando la barca, l'uomo ha inventato il naufragio, e scoprendo il fuoco ha assunto il rischio di provocare incendi mortali*».

Questa intrinseca rischiosità – tipica di un *pharmakos* – non deve ridurre la sua applicabilità, ma richiedere maggiori competenze, soprattutto **competenze che vadano oltre la sfera tecnica**. Inoltre, il fatto che l'IA sia un insieme di tecnologie recenti fa sì che il racconto dell'IA sia impastato di opinioni e manipolazioni, cosa che rende più difficile una comprensione e valutazione obiettiva del fenomeno.

Usando la illuminante espressione coniata da **Bruno Latour** per parlare sia di superstizioni religiose che di filosofia della scienza, potremmo dire che queste piattaforme sono un **"faticcio"**. Il termine – in originale *"faitiche"* – indica oggetti astratti costruiti dall'uomo per rappresentare fenomeni o fornire funzionalità. Il punto di partenza sono gli idoli prodotti dagli indigeni – oggetti artificiali che assumono proprietà e pertanto diventano "concreti" e vengono adorati come divinità esistenti. L'intuizione di Latour è che anche la scienza produce fatticci (ad esempio la catena del DNA, il neutrino, i buchi neri) per descrivere fenomeni naturali od oggetti tecnologici, che poi da iniziali e necessari strumenti concettuali assumono vita propria e diventano "reali".

Il fatticcio, dunque, nasce, per concettualizzare un fenomeno o una tecnologia innovativa... e per questo è utilissimo. Il rischio è che gradualmente, venga scambiato e poi trattato come fatto concreto. E ciò è particolarmente problematico quando gli interessi in gioco sono molteplici e compositi. Nel caso dell'IA, i fornitori, i finanziatori, i media alla ricerca compulsiva di novità, ma anche gli organismi di difesa delle Nazioni – che vedono nell'IA una grande minaccia/opportunità militare – rendono complessa la valutazione obiettiva di questo nascente gruppo di tecnologie.

Che fare dunque?

Non possiamo – né dobbiamo – fermare l'innovazione, anche quando è problematica e rischiosa. Anzi molte delle potenti innovazioni che fanno parte della nostra quotidianità hanno suscitato molto timore e pregiudizi al loro apparire.

L'importante è **aumentare la consapevolezza del loro modus operandi e finalità**, e soprattutto **sviluppare antidoti** per controllare le eventuali e insorgenti dimensioni problematiche e gli indesiderabili effetti collaterali.

L'aspetto più delicato dell'IA è, allora, che agisce sulla parte pregiata dell'essere umano: **il linguaggio, la conoscenza, le emozioni, la memoria**; interferisce, dunque, con il "pannello di controllo" dell'umano.

Concentriamoci allora su una specifica tecnologia che sta suscitando molto interesse: **l'IA generativa**. In questa fase dello sviluppo tecnologico non è possibile avere certezze, ma si può però suggerire un modo di osservare il fenomeno aperto e costruttivo ma critico. Anche in questo ambito, infatti, sta emergendo un pericoloso pensiero unico spinto dai produttori di queste tecnologie, dai finanziatori e ➤

“L’IA non indica una specifica tecnologia, quanto un gruppo di soluzioni la cui finalità è generare comportamenti autenticamente umani”

assecondato da un plotone di riviste e opinionisti sempre alla ricerca di sponsorizzazioni e di novità stupefacenti da sbattere in prima pagina.

Allora può farci da guida una riflessione del filosofo **Karl Popper**: «ogni qualvolta una teoria ti sembra essere l'unica possibile, prendilo come un segno che non hai capito né la teoria né il problema che si intendeva risolvere» (*Conoscenza oggettiva: un punto di vista evoluzionistico*, 1972). Anche la fretta di fare, capire e prevedere per normare e incentivare... va controllata il più possibile. **L'ansia di fare qualcosa per avere la sensazione di avere il controllo è molto pericolosa e spesso foriera di grandi errori.** La gatta che ha fretta fa i gattini ciechi, ci ricorda la saggezza delle nostre tradizioni.

Vi sono tre classi di utenti di questi sistemi: i **decisori**, gli **utilizzatori** e i **progettisti** che inseriscono questi sistemi nella quotidianità procedurale.

Mi concentrerò sui primi due; la terza tipologia di utente, infatti, richiede una trattazione a sé che sconfinava nell'ambito del software engineering e della progettazione organizzativa ed esula dalle finalità di questo articolo.

I decisori possono essere visti come dei super utenti che devono immaginarsi tutte le possibili applicazioni ed effetti di un sistema per valutarne pro e contro e deciderne l'adozione. Ma, a ben vedere, anche gli utilizzatori finali devono impraticarsi nell'arte della decisione in quanto devono valutare, di volta in volta, non solo quando usare una piattaforma di IA generativa.

Quali sono allora le competenze più critiche?

Tre in particolare:

- Sviluppare l'arte socratica di **fare le domande** – con il doppio obiettivo di ottenere maggiore ricchezza informativa e svelare eventuali manipolazioni – rafforzando il pensiero critico e attingendo alle tecniche dell'*agnotology*, la disciplina che studia come si «fabbrica» l'ignoranza, fondata recentemente da un professore di Stanford.

- Rafforzare le abilità di **infosourcing** imparando a collegare le risposte dell'IA con valutazioni verosimili sulla qualità e significatività delle fonti utilizzate e a variare l'uso delle fonti (riducendo l'eccesso di falsi positivi, neutralizzando le manipolazioni evidenti e cambiando il peso dei “segnali deboli”) per comprendere la “sensitività” dei risultati alle fonti (e cioè al loro peso).

- Prendere dimestichezza con il **funzionamento algoritmico dei computer** (chiamato tecnicamente *macchina di Von Neumann*) non tanto per imparare a programmare quanto piuttosto per comprendere le specifiche logiche di funzionamento dei computer (ad es. gli algoritmi ricorsivi o le varie tecniche di *indexing*) e le relative idiosincrasie (ad es. l'ambiguità o le definizioni circolari), e coglierne rigidità e limiti... un po' come fare pratica con le “tre leggi della robotica” di **Isaac Asimov**.

Inoltre, dobbiamo usare l'IA con efficacia e sistematicità, trattandola però come **allenatore e pre-fabbricatore veloce di semilavorati**, ma mai come sostituto. Anche perché, al contrario delle macchine che seguono pedissequamente le regole, i veri esperti si rendono conto quando sta emergendo un "cigno nero" nel loro dominio – un fatto cioè che smentisce la conoscenza consolidata – e allora sostituiscono la loro esperienza con l'intuito.

Un ultimo suggerimento: proteggersi dalla *captatio benevolentiae* digitale

È inoltre sempre più evidente che la finalità ultima delle piattaforme d'IA generativa non è tanto la produzione di conoscenza, di sistemi esperti, quanto la **realizzazione di un linguaggio potentissimo – detto language model – capace di adattarsi ad ogni interlocutore che le interroga e di persuaderlo della correttezza delle sue risposte**. Potremmo considerarlo una sorta di "certificato oracolare" che fa sì che queste tecnologie possano persuadere senza dover dimostrare né la coerenza di ciò che dicono né l'attendibilità delle fonti a cui attingono, che infatti rimangono sempre nascoste.

Per questo motivo il cuore del sistema è il *language model* e gli obiettivi del cosiddetto "allenamento" sono **dare in pasto alla piattaforma miliardi di frammenti di conversazione online** per allinearsi e diventare familiare con i linguaggi della Rete e con ciò che piace (e genera "like") e, in ultima istanza, rafforzare il proprio armamentario linguistico in senso persuasivo per arrivare a costruire ciò che **Cicerone** chiamerebbe *captatio benevolentiae* digitale.

Ritorna ancora la **centralità del pensiero critico e delle competenze linguistiche e retoriche** che ci consentono di smascherare le manipolazioni linguistiche, evidenziare i buchi logici o i ragionamenti fallaci, investigare la qualità delle

fonti e i livelli di attendibilità delle risposte.

L'Intelligenza Artificiale, quindi, rimette al centro la **cultura umanistica** nel senso più autentico, e cioè i saperi attorno all'essere umano, ciò che dobbiamo sapere per conoscerci. E uno degli aspetti critici di questa conoscenza sarà proprio il **nesso tra linguaggio e memoria**.

Può essere allora utile ai nostri ragionamenti una riflessione sulle memorie artificiali di **Boncompagno da Signa**, proveniente dal suo *Rhetorica novissima* del 1235 (ancora una volta un testo antico, per ricordarci che la comprensione dell'innovazione richiede anche una conoscenza del passato): «*Che cos'è la memoria. Memoria è un glorioso e ammirabile dono di natura, per mezzo del quale rievochiamo le cose passate, abbracciamo le presenti e contempliamo le future, grazie alla loro somiglianza con le passate. Che cos'è memoria naturale. Memoria naturale viene solamente da dono di natura, senza aiuto di alcuno artificio. Che cos'è memoria artificiale. Memoria artificiale è ausiliaria e assistente di memoria naturale [...] ed è chiamata "artificiale" da "arte", perché è stata scoperta "artificialmente", grazie a finezza d'ingegno*».

Pertanto, un utilizzo adeguato dell'IA richiede non solo competenze ma anche abilità di ragionamento e forza una maggiore comprensione di ciò che sappiamo, che è indissolubilmente legata a ciò che ricordiamo.

Ce lo ricorda **Dante** nella *Divina Commedia*: comprendere o imparare qualcosa non forma la vera conoscenza se non si tiene poi a mente ciò che si è imparato ("Non fa scienza senza lo ritenere, avere inteso"). E lo afferma in modo perentorio **Giuseppe Ungaretti** durante una conferenza su **Giambattista Vico**: «*Tutto, tutto, tutto, è memoria*».