

Tra Bellezza e Pericolo: le trappole delle intelligenze artificiali nell'arte

Gianluca Garofalo

C.atWork Creative

Alla ricerca di una perfezione estetica a basso costo, il pericolo dell'omologazione e del condizionamento è concreto. Le intelligenze artificiali generano capolavori digitali o insidie visive? Quali sono le trappole che si celano dietro la creatività algoritmica?

Il professore Paolo Benanti¹, in uno dei suoi numerosi interventi sul tema delle intelligenze artificiali (AI), osserva che la lente convessa è stata la causa di una enorme rivoluzione, che ha visto nascere il telescopio e il microscopio, oltre che aver spostato l'Uomo dal centro dell'universo ad una posizione periferica e minimale. Analogamente, l'avvento delle intelligenze artificiali ci obbliga ad osservare la realtà attraverso un *macroscopio*. Non è più l'immensamente grande o l'immensamente piccolo a spostare l'Uomo dalla sua centralità, ma l'immensamente complesso.

Osserviamo ed utilizziamo una tecnica che opera attraverso calcoli complessi su una mole incommensurabile di dati. Il problema, tuttavia, è che non siamo affatto preparati a farlo. Siamo, quindi, davanti ad un nuovo scenario di tipo antropologico, oltre che sociale.

In questo articolo cercheremo di fare chiarezza, procederemo per punti cardine, costruiremo dei piccoli fari che ci aiuteranno a comprendere meglio cosa abbiamo di fronte. Cercheremo, prima di tutto, di capire cosa siano questi sistemi di calcolo che chiamiamo intelligenza artificiale.

Le intelligenze artificiali non sono intelligenti

I sistemi di calcolo in esame sono estremamente performanti nel computare un'impressionante mole di dati, molto più di quanto potrebbe qualunque essere

¹ Presidente della Commissione AI per l'informazione, docente presso la Pontificia Università Gregoriana, l'Istituto Teologico di Assisi e il Pontificio Collegio Leoniano ad Anagni. Oltre ai corsi istituzionali di morale sessuale e bioetica, si occupa di neuroetica, etica delle tecnologie, intelligenza artificiale e postumano. Ha fatto parte della task force Intelligenza Artificiale che coadiuva l'Agenzia per l'Italia digitale. È membro corrispondente della Pontificia Accademia per la Vita con particolare mandato per il mondo delle intelligenze artificiali. A fine 2018 è stato selezionato dal Ministero dello Sviluppo Economico tra i trenta esperti che hanno il compito di elaborare le strategie nazionali sull'intelligenza artificiale e in materia di tecnologie basate su registri condivisi e blockchain.

umano, ma non ne hanno nessuna coscienza. Non sono provvisti di discernimento. Non hanno la capacità di comprendere il senso dei calcoli che eseguono. La tesi del filosofo Luciano Floridi² è che si sia operato un divorzio tra l'abilità di agire e l'intelligenza necessaria per agire. In altre parole, pur senza cognizione di causa, attraverso un calcolo statistico tra miliardi di immagini, testi o dati, il sistema riesce a proporre una risposta alla nostra interrogazione sufficientemente soddisfacente. Questo perché la risposta è in linea con la media che consideriamo accettabile. L'impressione è che risponda in modo senziente, ma è solo un'impressione, appunto.

La narrazione legata al fenomeno è fuorviante

Utilizzare termini come *imparare*, *apprendere*, *decidere*, *addestrare*, altera la percezione reale del fenomeno. In effetti, nel momento in cui si prova a chiamare il *machine learning*, ovvero l'addestramento dell'intelligenza artificiale, con il suo vero nome, cioè calibrazione statistica del sistema di calcolo, la poesia svanisce.

Il fatto è che questi sistemi non imparano affatto, ma subiscono una regolazione per aumentare la precisione del calcolo, oltre ad un incremento dei dati a disposizione. Questi dati, tra l'altro, sono spesso coperti da copyright, ma tale realtà sembra assumere l'aspetto di un piccolo particolare irrilevante, soprattutto agli occhi di chi sostiene una visione dello sviluppo tecnologico che può essere definita religiosa, anche e, forse, principalmente, nell'espressione artistica e creativa dell'Uomo. Adottando la definizione proposta da Erich Fromm, la religione è qui intesa in senso lato, come «un sistema di pensiero e azione condiviso da un gruppo, che offre all'individuo un mezzo di orientamento e un oggetto di devozione»³.

Etichettatura o sfruttamento?

La macchina, ad un certo punto, senza un input specifico fornito dal tecnico, inizia a computare autonomamente l'enorme massa di dati provenienti da tutto ciò che, in qualche modo, è online. Il meccanismo di raccolta ed etichettatura può essere automatico, attraverso dei bot che raccolgono dati direttamente dalla rete, oppure manuale. I dati arrivano alla macchina solo dopo che un gran numero di persone, lavoratori del terzo mondo, il più delle volte sottopagati, ha "etichettato" i dati come conformi o meno. Proprio la necessità di rendere i dati conformi ad un criterio predefinito, sottopone queste persone ad un forte stress psicologico, per la necessità di escludere immagini non socialmente opportune come stupri, violenze, anche su minori, pornografia di ogni genere e tanto, tanto altro. Da un certo punto di vista, quanto accade non è poi così diverso dal colonialismo dei secoli tra-

² Professore ordinario di Filosofia ed etica dell'informazione presso l'Oxford Internet Institute della Oxford University, per la quale dirige dal 2017 il Digital Ethics Lab; chairman del Data Ethics Group dell'Alan Turing Institute.

³ Erich Fromm, *Avere o essere?*, Mondadori: Milano, 1976, p. 208.

scorsi, ovvero uno sfruttamento di lavoratori provenienti da paesi che hanno una scarsa capacità di tutela del lavoro. Niente di nuovo, purtroppo.

Da questo punto di vista, le scintillanti AI appaiono piuttosto opache.

Proprio questa necessità di conformità definita dall'azienda che produce il sistema, dà origine ad un pregiudizio molto condizionante, definito "pregiudizio algoritmico". Si tratta di una distorsione di partenza, una discriminazione che nasce prima ancora che i dati vengano raccolti e processati. Si tratta di un bias di tipo statistico. Se un dato è zero, non vuol dire che l'oggetto a cui fa riferimento non esista. Un esempio ampiamente riportato da diversi studiosi è quello legato alla previsione, da parte di una AI predittiva, di quale sarà il prossimo presidente degli Stati Uniti d'America. Il risultato ha dato indicazioni abbastanza precise: tendenzialmente di pelle bianca, sopra la sessantina, sicuramente uomo. La previsione è stata fatta sui dati del passato, sulle caratteristiche di tutti i presidenti eletti alla Casa Bianca. Dati acquisibili, non ipotetici, e non ci sono dati relativi a donne che abbiano ricoperto questo ruolo. In effetti, non è mai successo che una donna salisse alla Casa Bianca. Un altro esempio che spiega molto bene il pregiudizio algoritmico, è il caso della ragazza afgana. Utilizzando "ragazza afgana" nel prompt per ottenere un'immagine da un'AI generativa, il risultato è stato una quantità pressoché infinita di variazioni sul tema della famosa fotografia di Steve McCurry, l'ipnotica ragazza afgana dagli occhi verdi pubblicata nel giugno del 1985 sulla copertina del National Geographic. Tutte le ragazze afgane, per l'AI, hanno solo e soltanto quelle specifiche caratteristiche. Il dato trascritto con quel titolo coincide solo con quell'immagine, perché è talmente alto il grado di incidenza nel calcolo statistico che non ve ne possono essere altri. Ma noi sappiamo bene che di ragazze afgane ne esistono anche di molto diverse da quella raffigurata in quel meraviglioso ritratto.

Non solo. Si tratta di un problema che assume contorni preoccupanti anche in virtù di un altro gap, ossia quello delle persone senza competenze tecniche a cui è affidato il compito di etichettare i dati.

Il calcolo procede a livello sub-simbolico

Il sistema, dal momento in cui inizia a procedere autonomamente, opera un calcolo seguendo algoritmi che non sono più riconducibili ad una calibrazione di partenza e, quindi, non sono più così correggibili.

In altre parole, non siamo più in grado di collegare la reazione alla causa. Non sappiamo più comprendere perché la macchina dia determinate risposte, accettabili o inaccettabili che siano. Non abbiamo più il controllo.

Possiamo, a questo punto, considerare la macchina come una figura giuridica, che agisce autonomamente e di propria volontà? Possiamo sollevare l'umano, l'autore del prompt, dalla responsabilità di esserne l'artefice? A chi assegnare diritti e doveri? Il prompt designer è davvero l'autore dell'immagine e del testo generati? Chi è l'artista? L'Uomo decide o sceglie tra le decisioni della macchina?

Proprio in questo tipo di dinamiche si inserisce un fenomeno piuttosto noto: il feedback loop. È un fenomeno che si innesca nel momento in cui l'umano interagisce con l'AI. L'umano inserisce un input nel sistema e il sistema risponde. Fin qui sembra procedere tutto come ci si aspetterebbe. In seconda battuta, però, per calibrare il risultato, l'umano inserisce un nuovo input basato sulla precedente risposta della macchina, e la macchina risponde nuovamente. Dopo un certo numero di cicli è estremamente difficile capire se l'intenzione umana sia libera dal condizionamento del sistema stesso. In altre parole, non è più così facile capire chi decide, chi tiene il timone, se l'umano o la macchina. In questo loop non ha alcuna importanza il fatto che la macchina non abbia una volontà sua, il condizionamento è indipendente da questo, ma coerente con il bagaglio di dati che il sistema utilizza. Torniamo, quindi, ad essere condizionati da un pregiudizio algoritmico.

Da qui nasce il dubbio se uno strumento tecnico come questo sia davvero neutro e neutrale. Del resto, sappiamo bene che nessuna tecnica può definirsi neutrale.

La non neutralità della tecnica

Ecco un altro aspetto fondamentale per comprendere a fondo il motivo che rende ambiguo l'uso di questi sistemi. Scrive Martin Heidegger nel suo *La questione della tecnica*:

«Non possiamo quindi esperire veramente il nostro rapporto con l'essenza della tecnica finché ci limitiamo a rappresentarci la tecnicità e a praticarla, a rassegnarci ad essa o a fuggirla. Restiamo sempre prigionieri della tecnica e incatenati ad essa, sia che la accettiamo con entusiasmo, sia che la neghiamo con veemenza»⁴.

Hannah Arendt, nel suo *La banalità del male*, evidenzia come un'azione legata ad una narrazione efficace, anch'essa una tecnica, possa rendere accettabili e, addirittura, comportamenti virtuosi, le più terribili nefandezze. Nel testo, Arendt riporta le parole del gerarca nazista Adolf Eichmann:

«Hitler, disse, "avrà anche sbagliato su tutta la linea; ma una cosa è certa: fu un uomo capace di farsi strada e salire dal grado di caporale dell'esercito tedesco al rango di Führer di una nazione di quasi ottanta milioni di persone... Il suo successo bastò da solo a dimostrarmi che dovevo sottostargli". E in effetti la sua coscienza si tranquillizzò al vedere lo zelo con cui la "buona società" reagiva dappertutto allo stesso suo modo. Egli non ebbe bisogno di "chiudere gli orecchi", come si espresse il verdetto, "per non ascoltare la voce della coscienza": non perché non avesse una coscienza, ma perché la sua coscienza gli parlava con una "voce rispettabile", la voce della rispettabile società che lo circondava»⁵.

L'ermeneutica del manufatto tecnologico e la sua vera essenza evidenziano un da-

⁴ Martin Heidegger, *La questione della tecnica*, Firenze: GoWare, 2017.

⁵ Hannah Arendt, *La banalità del male*, Milano: Feltrinelli, 2023 (prima edizione digitale).

to essenziale: la tecnica risponde ad una domanda dell'Uomo. Questa domanda è ciò che l'Uomo immagina di dover avere per risolvere e migliorare la sua esistenza. In una condizione di benessere e fertilità intellettuale, l'Uomo inventa l'agorà, la filosofia e l'organizzazione sociale democratica ed etica. In una condizione di paura, malessere e pericolo, l'Uomo inventa la falange oplitica, le armi e la guerra. A quale tipo di società dobbiamo ascrivere l'invenzione delle AI?

Una considerazione interessante va fatta a riguardo delle AI che suggeriscono consigli sulle piattaforme di commercio online: «forse ti interessa anche quest'altro articolo...», «altre persone hanno acquistato anche...». Non è raro pensare che, effettivamente, la proposta sia interessante, ma ne abbiamo davvero bisogno? Avremmo scelto proprio quell'articolo se non avessimo avuto quel "consiglio"? Sembra un servizio su misura, ma a ben vedere è un'opera di omologazione.

La rivoluzione big data

Dopo aver inquadrato cosa siano i sistemi definiti Artificial Intelligence, cerchiamo ora di entrare un po' di più nella questione sociale e antropologica.

Sappiamo che l'Uomo nasce tecnologico, come specie.

Con il termine tecnica intendiamo cose, sistemi, anche molto diversi tra loro, ma che hanno tutti la stessa caratteristica: rispondere ad una domanda, ad una necessità dell'Uomo.

Così, per una domanda di comunicazione molto pressante, nascono un codice, un simbolo, un'icona e un linguaggio. Seguono sistemi per esprimere il senso, per trasportare le idee, per lasciare istruzioni, per superare l'assenza e indicare dei modi. Segno, disegno, scrittura, fino al nostro contemporaneo che ci consegna un fatto sorprendente: solo tra il 2013 e il 2015, l'umanità ha prodotto più dati che in tutta la sua storia⁶ ed è stata in grado di trasmetterli velocemente come mai prima. Come se tutta la produzione, dall'origine dell'umanità fino al 2012, tra romanzi, saggi, poesia, canzoni, musica, scienza, pittura, scultura, architettura, fotografia, cinema, tecnologia, non bastasse ad eguagliare la quantità di opere prodotte in soli due anni. Come se in questi due anni, per fare un paragone di tipo temporale, avessimo visto nascere il Partenone qualche ora prima del lancio dell'Apollo 11.

Questa enorme quantità di dati aumenta quotidianamente. Si è stimato che nel 2020 abbiamo prodotto circa 44 zettabyte di dati e li abbiamo fatti circolare istantaneamente per il globo, ovvero 440 miliardi di gigabyte alla velocità della luce⁷.

Com'è possibile?

In realtà, comprendere come questo sia avvenuto è abbastanza semplice: date uno

⁶ Bernard Marr, *The Complete Beginner's Guide to Big Data Everyone Can Understand*, «Forbes», (14 marzo 2017), <<https://www.forbes.com/sites/bernardmarr/2017/03/14/the-complete-beginners-guide-to-big-data-in-2017/>>.

⁷ Federica Brancale, *Big Data: Cosa sono, la storia, le caratteristiche, le analisi ed esempi*, Marketing Freaks, 2024, <<https://www.themarketingfreaks.com/2019/11/big-data-cosa-sono-la-storia-le-caratteristiche-le-analisi-ed-esempi/>>.

smartphone a tutti e inventate dei luoghi virtuali dove riversare parole, foto, video, disegni, citazioni, news, fake news, stati d'animo, diari personali, giochini di riconoscimento facciale, test psicologici, comizi, opinioni...

In due anni il gioco è fatto!

Il primo smartphone realmente tale, cioè come lo immaginiamo oggi, fa il suo esordio tra il 2008 e il 2009. Facebook, ciò che possiamo definire il social network più famoso e usato al mondo (stacca ancora di poco YouTube), l'erede dei forum, nato nel 2004 come realtà legata alla vita universitaria di Harvard, poi Stanford, Columbia University e Yale, Boston University e Boston College, proprio nel 2009, uscito dai confini dei campus, registra il suo primo anno in attivo. Solo in Italia, nell'agosto del 2008, Facebook registra un incremento di iscrizioni del 961% rispetto al mese precedente.

La domanda dell'Uomo

Solo un paio di decenni fa, quest'enorme quantità di dati avrebbe chiesto macchine estremamente costose, di grandi dimensioni, le risorse computazionali non sarebbero state adeguate, e avrebbe avuto bisogno di tanto, tanto tempo per essere analizzata, utilizzata e trasformata. Oggi possiamo farlo con un portatile, un tablet o uno smartphone, grazie all'innesto degli algoritmi capaci di trattare così tante variabili in poco tempo.

Possiamo digitare un prompt sulla tastiera e ritrovarci una splendida immagine sul monitor. Con due click mostrare al mondo il piatto che stiamo degustando in un ristorante nella periferia di Milano. Facendolo, produciamo dati ad una velocità assurda.

Ma possiamo davvero immaginare che la qualità delle pubblicazioni di Dante Alighiero, sul suo profilo Facebook, sia equiparabile alla Divina Commedia? O che il selfie di Maia Lino abbia un valore culturale degno di nota quanto l'autoritratto con passaporto ebreo di Felix Nussbaum?

Nel 2005 David Foster Wallace, profeticamente, pronunciò, davanti alla classe di laureati del Kenyon College, un discorso che in seguito diventò molto famoso e che, con una storiella spiritosa, individuava molto bene una caratteristica essenziale della nostra specie.

«Ci sono questi due giovani pesci che nuotano e incontrano per caso un pesce più grande che nuota dall'altra parte, che fa loro un cenno e dice: "Buongiorno, ragazzi. Com'è l'acqua?". I due giovani pesci nuotano per un po' e alla fine uno di loro guarda l'altro e dice: "Che diavolo è l'acqua?"»⁸.

⁸ David Foster Wallace, *Questa è l'acqua*, Torino: Einaudi, 2009. Così il grande scrittore americano introduceva il suo discorso di congratulazioni ai laureati del prestigioso Kenyon College nel 2005; un intervento divenuto famoso con il titolo "This is water" in cui l'autore di *Infinite Jest* spiegava ai laureandi il vero valore dell'educazione superiore, quella capacità di analizzare l'essenza della realtà che ci circonda così da costruire delle possibilità di scelta.

L'Uomo è una specie estremamente adattabile e questo può essere un pregio, ma anche un difetto. Dopo un certo tempo si abitua e smette di avvertire come estraneo, estraneo, qualunque dispositivo tecnico stia adoperando. Occhiali, magliette, calzini, linguaggio, penne, pennello, computer, smartphone, algoritmo.

Di fatto, questo è un processo di incorporazione che sfoca la visione della realtà sempre di più, via via che la distanza tra la mano e il *manicum* tecnologico aumenta. Via via che la tecnologia diventa più complessa e meno immediatamente riconoscibile e indagabile.

Sarà capitato a tutti di eseguire alcuni gesti automatici che, però, ci lasciano confusi e che hanno origine nel mondo digitale. Cercare di zoomare l'immagine di un libro, scrollare una fotografia per vedere se si riesce a scorgere altri dettagli sopra o sotto la cornice che la definisce, pensare di poter tornare indietro quando si sbaglia qualcosa, si lascia cadere un bicchiere, premendo contemporaneamente Control+Z.

Quello che abbiamo creato è un ambiente vitale a logica digitale e guidato interamente a livello digitale. Un ambiente che ci condiziona anche dove il nostro agire dovrebbe essere analogico, per così dire. Il nostro pensare è fortemente condizionato da questa logica virtuale.

Tutta questa massa di dati, di immagini, di opinioni, di post, che riversiamo nel web ogni minuto, da circa dieci anni, è davvero, come sostiene qualcuno, uno specchio, un'immagine fedele della nostra umanità? Della nostra società? Questi atteggiamenti sono davvero liberi?

Quando diamo alle AI generative di immagini il ruolo di *veicoli di democrazia nell'arte*, siamo davvero coscienti dell'azione che stiamo supportando? Questa visione postumanista e transumanista che vede la salvezza dell'umano, la sua liberazione, attraverso l'azione della tecnica, verso uno stato ideale di nuovo Eden, di «paradiso disincarnato del cyberspazio»⁹, è davvero l'evoluzione più auspicabile, la più giusta?

Possiamo sospendere ogni giudizio etico e accettare l'innovazione come progresso, senza nessuna azione? Passivamente?

Sono convinto che senza una opportuna educazione e senza un limite normativo umano, rischiamo di non essere più in grado di utilizzare e vivere l'ambiente digitale con consapevolezza e intenzionalità.

Il governo della tecnica sull'Uomo e non viceversa, si attua attraverso questo meccanismo di incorporazione. Il Control+Z da utile strumento si trasforma in *forma mentis*. L'azione non è più quella di un uso consapevole, di un governo umano dell'utente, ma, al contrario, un intervento della macchina a livello percettivo, interpretativo, neuronale.

⁹ Hava Tirosh-Samuels, *Transhumanism as a secularist faith*, «Zygon: Journal of Religion and Science», 47 (2012), n. 4, p. 710-733.

Allo stesso modo, la scelta di immagini generate da prompt linguistici, apparentemente governata dall'operatore, in realtà propone scelte determinate, condizionate dal sistema selettivo stesso della macchina, dall'azione dell'algoritmo.

L'azione arbitraria del calcolo avviene indipendentemente dal bagaglio esperienziale umano dell'operatore. Le *reference*, lo spazio culturale, emotivo, diventano aziendali e non personali.

Di fatto, potremmo trovarci davanti ad una sostituzione condizionata e arbitraria della cultura visiva individuale.

Per questi motivi, non credo sia corretto demonizzare una tecnologia innovativa come i sistemi di AI, ma, allo stesso modo, non credo sia opportuno porsi con un approccio religioso, mitologico, al tema. È imprescindibile vedere la natura ambivalente dello sviluppo tecnico, che non è mai scontato nell'imprevedibilità delle metodologie ed è sempre figlio della realtà politica, sociale e culturale contingente.

Bypassare le azioni umane di progettazione e, conseguentemente, di comunicazione, o limitarle a una scelta condizionata che mima un'arbitrarietà, ma che esclude la radice stessa dell'azione percettiva, interpretativa e reattiva del cervello umano all'origine, trasforma nella sostanza e non solo nella forma un sistema di comunicazione e di educazione, culturale e quindi sostanziale, tipico e peculiare della nostra specie.

Di fatto, abdichiamo alla nostra intelligenza, non soltanto alla nostra abilità di fare. Qual è la domanda alla quale questa tecnica risponde?

In the pursuit of low-cost aesthetic perfection, the danger of standardization and conditioning is real. Do artificial intelligences generate digital masterpieces or visual pitfalls? What are the traps hidden behind algorithmic creativity?

L'ultima consultazione dei siti web è avvenuta nel mese di giugno 2024