



A JOURNAL OF THE
SOCIAL IMAGINARY

Bambini artificiali. La metafora del bambino e l'immaginario dell'IA

Paolo Bory

paolo.bory@polimi.it

Dipartimento di Design | Politecnico di Milano



Abstract

"Artificial children". The child-metaphor and the AI imaginary

This article explores the pervasive metaphor of anthropomorphizing artificial intelligence (AI) as a child, tracing its historical roots and examining its current implications. Despite warnings against such anthropomorphizing due to ethical and practical risks, the child-like metaphor persists in AI discourse, both historically in foundational texts by figures like Alan Turing and Norbert Wiener, and in contemporary big tech and media narratives. The metaphor influences public perception, reinforcing ideas of AI's innocence and potential, while often sidelining the complexities and impacts of current AI technologies. The article critiques this view, highlighting the risks of underestimating AI's present capabilities and the importance of reframing the imaginary of AI towards a responsible integration of this technology in our everyday life.

Keywords

AI | Metaphors | Narratives of AI | Imaginary | Children



Introduzione

Una delle regole auree che circola maggiormente tra giornalisti e accademici che studiano o coprono il campo dell'intelligenza artificiale sostiene che non si dovrebbe mai "antropomorfizzare l'IA" (Heikkilä, 2023). Raccontare, descrivere, ricostruire il funzionamento dell'IA utilizzando l'essere umano come analogia o specchio è secondo molti studiosi e studiose sbagliato e rischioso, in quanto comporta diversi dilemmi e risvolti di natura etica, di mistificazione, e in alcuni casi forme di discriminazione e incoerenza sia teorica che empirica con la realtà dei fatti (Fossa 2021; Placani, 2024). Eppure, la storia dell'IA, così come quella dei computer, è a partire dai suoi albori caratterizzata dal parallelismo con l'essere umano, sia nella sua completezza che con il suo intelletto; i cervelloni o cervelli elettronici, le reti neurali, i cyborg, gli androidi, i robot, e ancor prima gli automi sono segni inequivocabili di come l'IA sia un prodotto che insiste da sempre sul principio prometeico e creativo del "a nostra immagine e somiglianza" (Natale 2021).

Ancora oggi sono innumerevoli le strategie retoriche, iconografiche e narrative che veicolano questo immaginario speculare, che va ben oltre la letteratura e la produzione fantastica e fantascientifica (Bory & Bory, 2015). Da un punto di vista iconografico, basta pensare a quanti eventi, organizzazioni, seminari e conferenze sull'IA vengono pubblicizzati con l'immagine della creazione di Michelangelo in cui al dito di Dio si sostituisce quello dell'uomo, e a quello di Adamo il dito robotico (Fig. 1).

Nonostante l'imperativo di non antropomorfizzare l'IA, sembra quasi impossibile, e paradossalmente non umano, sfuggire a questa tentazione. Da un lato la relazione e soprattutto la familiarizzazione con la tecnica, come dimostrano gli studi sugli immaginari e sui processi domesticazione, necessitano spesso di ponti che permettono euristicamente l'inclusione delle nuove tecnologie nel quotidiano, compresa quelle interazioni apparentemente innocue e innocenti con e attraverso di esse. Allo stesso tempo, se è vero che il parallelo IA-umanità è fallace, alla critica di questo principio di specularità gli studi sociali, filosofici e antropologici sull'IA non possono esimersi dal decifrare le implicazioni culturali, economiche, politiche e anche squisitamente pratiche e di vita quotidiana che determinate rappresentazioni della tecnologia veicolano, incluse le metafore e le analogie che gli attori istituzionali, economici e scientifici, così come gli utenti, adottano per raccontare queste tecnologie e costruire degli specifici immaginari a partire e su di esse (Bucher 2019; Mager & Katzenbach, 2021). Le metafore hanno quindi un ruolo cruciale nel processo di domesticazione dell'IA, come sostengono Lakoff e Jhonson:

Metaphors have entailments through which they highlight and make coherent certain aspects of our experience. [...] A metaphor may thus be a guide for future action. Such actions will, of course, fit the metaphor. This will, in turn, reinforce the power of the metaphor to make experience coherent. In this sense metaphors can be self-fulfilling prophecies. (Lakoff e Johnson, 2008: 157)





Le metafore hanno da sempre rivestito la tecnologia di significati che richiamano elementi naturali, come nel caso della rete o del cloud per Internet (Hu, 2015; Peters, 2015, Wyatt, 2021) così come nel campo dell'intelligenza artificiale (Cave et. Al 2020; Hermann 2023), ponendosi a volte come intermediazioni ideologiche, come immagini utopiche e distopiche dei futuri possibili, o anche come forme ingannevoli e fin troppo semplificatorie dei complessi apparati materiali e umani che creano, promuovono e mantengono vive le infrastrutture digitali (Bory, 2020). Ma le metafore hanno anche il potere di orientare l'immaginario verso un preciso percorso di stabilizzazione e di istituzionalizzazione (Castoriadis, 1975) della tecnica, radicando non solo una narrazione condivisa ma anche un suo dispiegamento e un ruolo sociale, economico e politico preciso. Nonostante gli immaginari sociali dell'IA, come ha giustamente sottolineato Taina Bucher (2019), non corrispondano per forza di cose agli immaginari individuali così come costruiti dagli utenti nei processi di interazione con le tecnologie algoritmiche, l'immaginario sociotecnico dell'intelligenza artificiale è in una fase profondamente magmatica (Castoriadis, 1975), un periodo cruciale e pre-politico (Mosco, 2005) di costruzione e ri-costruzione, in cui gli usi particolari e il racconto dell'IA nella sfera pubblica si nutrono, si scontrano, mutano vicendevolmente.

Aldilà delle visioni utopiche delle Big Tech e dei tecnoentusiasti, allo stato attuale permane una profonda incertezza, se non una vera e propria contraddizione da sciogliere criticamente, tra l'immaginario inteso come previsione, rappresentazione e proiezione della tecnologia nel prossimo futuro, e l'identità, l'uso pragmatico e fattuale dell'IA allo stato attuale.

Dall'avvio di questa nuova "estate" dell'IA, che storicamente si alterna a momenti di ibernazione o "inverni" in cui il termine scompare dal dibattito tecno-scientifico, l'immaginario sociotecnico si è soffermato maggiormente sulla trasformazione di questa tecnologia in una *super* o *general intelligence* (Bostrom 2014), anche considerando i rischi che l'avvento di una intelligenza superiore a quella umana porterebbe con sé, prima di tutto per la sopravvivenza stessa della nostra specie.

Questa narrazione permea ancor prima che la società civile gli stessi luoghi dove l'IA viene progettata e prodotta, come ha raccontato Steven Levy (2023) nel suo articolo sull'azienda leader del mercato Open AI. L'obiettivo prometeico degli impiegati di queste aziende sembra essere proprio quello di creare la superintelligenza, di raggiungere quindi il punto alpha in cui la singolarità prenderà forma. Ma in realtà si tratta solo di uno dei modi in cui l'IA viene descritta e immaginata, soprattutto quando si parla del tipo di tecnologia ancora limitata e controllabile che adottiamo adesso.

Quasi, anche se solo apparentemente, in contraddizione con l'idea di un'intelligenza superiore, il modo di raccontare lo sviluppo attuale nonché l'identità di questa tecnologia che chiamiamo IA, o di questo insieme di tecnologie se vogliamo essere più precisi, va di pari passo con una metafora particolarmente efficace: quella del bambino o della ragazzina, dell'infante che deve essere cresciuta, curata, ed educata *oggi* per sfruttarne al massimo le potenzialità *domani*, in modo tale che i suoi comportamenti e le sue azioni siano orientati unicamente al bene dell'umanità.

Questa metafora per immaginare, pensare e descrivere l'IA non è nuova, né realmente in contraddizione con l'immaginario della *super intelligence* (Goode, 2018).

Come cercheremo di mostrare, si tratta piuttosto di un ponte semantico estremamente conveniente per favorire l'accettazione sociale della tecnologia, spostando anche l'attenzione da temi urgenti relativi al suo impatto quotidiano, favorendo la reiterazione di un immaginario che vuole essere "imposto" dagli attori dominanti del mercato digitale, plasmando, o anche spezzando quei fili della vita quotidiana che, come diceva Simmel, intessono le trame della vita collettiva (Marzo, Meo, 2013: 8).

In questa sede, affronteremo alcuni dei significati e dei rischi derivanti da questa metafora dell'IA bambina, le sue implicazioni sociali e culturali, nonché le sue dinamiche politiche, seguendo tre passaggi analitici. In primo luogo, tratteremo alcuni precedenti storici particolarmente significativi nell'uso di questa metafora in testi e autori chiave come nel caso del famoso articolo di Alan Turing sul rapporto tra meccanica artificiale e intelligenza umana. Nella sezione successiva, alcuni esempi emblematici del contemporaneo come il caso delle rivelazioni dell'ingegnere di Google Blake Lemoine, secondo cui il modello di linguaggio LaMDA è senziente, verranno analizzati per approfondire il ruolo dell'analogia IA/bambino al giorno d'oggi. Infine, la sezione finale dell'articolo intende rovesciare il ruolo paternalistico nei confronti dell'IA mettendo in evidenza il rischio di infantilizzazione in cui gli utenti in particolare incorrono attraverso le dinamiche ludiche che spesso non solo sminuiscono il ruolo delle IA di oggi sul nostro quotidiano, ma che compromettono anche il significato e il valore del nostro contributo come utenti alla sua diffusione, nonché alla sua costruzione e accettazione in ambito sociotecnico e culturale.



FIG. 1 – Uno dei tanti esempi di imitazione della creazione di Michelangelo, e allo stesso tempo un raro esempio di una mano femminile invece che maschile.

(Fonte: Winstead 2021)

1. I padri dell'IA

Se andiamo indietro nel tempo, l'analogia dell'IA bambina non è affatto nuova, ma è una costante nella storia dell'informatica e della cibernetica, sin dai loro albori. Già Alan Turing nel suo articolo fondativo pubblicato da *Mind* del 1950 scriveva:

Invece di elaborare un programma per la simulazione di una mente adulta, perché non proviamo piuttosto a realizzarne uno che simuli quella di un bambino? Se la macchina fosse poi sottoposta ad un appropriato corso d'istruzione, si otterrebbe un cervello adulto. Presumibilmente il cervello infantile è qualcosa di simile ad un taccuino di quelli che si comprano dai cartolai. Poco meccanismo e una quantità di fogli bianchi [...] La nostra speranza è che ci sia così poco meccanismo nel cervello infantile, che qualcosa di analogo possa venir facilmente programmato. (Turing, 1950: 438).



Da questo stralcio, con gli occhi di oggi, sembra quasi che una delle menti più brillanti della storia del Novecento si sia esposta a una facile critica. Ma anche a prescindere dall'aspetto tecnico e teorico di base per cui l'IA è come un bambino/pagina bianca da riempire, è evidente che il padre dell'intelligenza artificiale non è mai stato realmente padre. Considerare la mente dei bambini come una risma di fogli bianchi esclude senza alcuna razionalità tutti quegli elementi costitutivi, biologici, ambientali e psicologici dell'apprendimento, nonché l'elaborazione non solo razionale, ma la pluralità di relazioni, la componente emozionale e quella puramente corporea e fisica del mondo che li circonda.

In un breve lavoro critico sul pensiero di Turing la studiosa Teresa Heffernan ha decostruito e messo in discussione gran parte delle premesse teoriche, epistemologiche e addirittura teologiche di Turing soprattutto quando egli sostiene che:

Tentando di costruire macchine del genere non vogliamo usurpare irriverentemente il Suo potere di creare anime, più di quanto non facciamo procreando bambini: piuttosto noi siamo, in entrambi i casi, strumenti del suo volere, fornendo abitazioni per le anime che Egli crea. (ibidem, 1950: 443).

Seguendo questo pensiero escatologico, secondo Turing saremmo quindi "figli di Dio" che procreano i loro figli meccanici. Ma aldilà della discutibilità del ruolo e la relazione con il divino, comprensibile da un punto di vista storico, ciò che anche Heffernan e studiosi come Katherine Hayles sottolineano è l'assenza ricorrente di "madri" nell'immaginario dell'IA; ciò rispecchia da un lato la dominanza maschile e l'attitudine paternalistica che abbiamo da sempre verso le "macchine pensanti", e dall'altro l'assenza di una consapevolezza, o di una volontà di racconto, del ruolo delle donne nella storia dei calcolatori, e anche di alcune differenze profonde tra stili di pensiero nell'immaginario sociotecnico delle macchine pensanti, come il ruolo della creatività e della poetica che contraddistingueva il percorso biografico e teorico di altre figure mitologiche della storiografia digitale come Ada Lovelace.

Un altro *padre* dell'IA come Norbert Wiener adottava un'analogia simile a quella di Turing per descrivere il processo pedagogico di "formazione" delle macchine.
Scriveva Wiener nel suo *Dio e il Golem*:

Mi sembra che la migliore speranza di una traduzione meccanica ragionevolmente soddisfacente sia quella di sostituire un meccanismo puro, almeno in un primo momento, con un sistema meccanico-umano, che coinvolga come parte critica un traduttore umano esperto, il quale insegna con degli esercizi *come un insegnante istruisce i suoi scolari*. (Wiener, 1996: 79, traduzione e corsivo dell'autore)

Sulla stessa onda, diversi esperti di oggi reiterano questo *topos* dello scolaro da educare o del bambino da crescere. Fang Chen, professoressa dell'University of Technology di Sydney, nota per i suoi studi sull'IA, ha sostenuto in un'intervista che:

L'intelligenza artificiale è come un bambino [...] Noi insegniamo a lei o al sistema a fare qualcosa. Le influenziamo, diamo loro alcuni principi, e poi a seconda di come lo progettiamo, il sistema va avanti. (Chen, 2018, traduzione dell'autore)

In molte conferenze e presentazioni di ingegneri, artisti e designer, questa metafora del bambino ritorna come un mantra, come uno strumento discorsivo di cui difficilmente riusciamo a fare a meno per parlare dell'IA di oggi, ma non solo. Riviste e giornali specializzati negli ultimi anni hanno fatto emergere la rilevanza di questa metafora nelle ricerche e nelle sperimentazioni scientifiche sull'IA.

L'MIT ha lanciato un'iniziativa per sviluppare un'intelligenza artificiale che impara come i bambini (Thys & Lepiarz, 2018)

La prima bambina AI è stata svelata a Pechino da uno scienziato cinese che ha lasciato gli Stati Uniti per la Cina (Tong, 2024)

È nata un'IA che pensa e impara proprio come un bambino di tre mesi (Daly, 2022)

Questa intelligenza artificiale impara come un bambino (Intagliata 2022)

Scienziati hanno istruito un'IA attraverso gli occhi di un bambino (The Economist, 2024)¹

Si tratta solo di alcuni dei tanti titoli che raccontano il progresso dell'IA non solo metaforicamente, ma anche pragmaticamente come un percorso di apprendimento infantile. Nel settembre del 2020, molto prima del successo di Chat-GPT il *Guardian* ha fatto scrivere a GPT-3 di Open AI un articolo intitolato "A robot wrote this entire article. Are you scared yet human?" (GPT-3 2020) nel quale l'IA fondamentalmente cercava di dimostrare la sua utilità, il suo istinto di sopravvivenza e le sue intenzioni

¹ Tutti i titoli sono stati tradotti dall'inglese dall'autore.



non belligeranti nei confronti dell'umanità. Come molte altre conversazioni e articoli che gli sono seguiti negli ultimi anni, si tratta di un'opera poco credibile, e, come molti altri autori hanno sostenuto, per molti versi banale e scontata, paragonata a un breve tema di scuola scritto, ancora una volta, da una ragazzina o un ragazzino.

Risulta chiaro da questi esempi che il tentativo di antropomorfizzare l'IA non è solo trasversale all'ambito mediatico e scientifico, ma è quasi automatico e dato per scontato da molti attori in gioco, e la metafora di Turing ha avuto conseguenze anche molto rilevanti sul lungo corso dell'IA fino a tempi recentissimi.

2. Innocenza e innocuità

Ci sono diversi problemi nel definire l'IA come una bambina, che vanno di pari passo con altrettanti rischi. In primo luogo, come accennato, stiamo assistendo a una dilagante, quanto preoccupante, forma di paternalismo fortemente declinata al maschile. Ne sono testimonianza le tante posizioni preventive prese dai CEO delle big tech, come nella controversa lettera firmata da nomi come Elon Musk e Steve Wozniak in cui si chiede di fermare le sperimentazioni sui modelli di linguaggio naturale (Future of Life Institute 2023). Una lettera da cui si evince la paura della crescita di una creatura, come quella di Frankenstein, potenzialmente incontrollabile, mentre dall'altro lato traspare l'assunto per cui solo dei "buoni padri", preferibilmente maschi, bianchi e statunitensi, possono dare la giusta *disciplina* alle macchine pensanti per salvare il nostro futuro.

Contemporaneamente l'immagine del bambino, del figlio, della ragazzina, nutrono un grande vantaggio comunicativo perché veicolano inevitabilmente un senso di innocenza e di innocuità; come se l'IA, in quanto entità innocente, fosse solo *potenzialmente* pericolosa, ma allo stato attuale delle cose del tutto innocua.

Eppure, i danni dell'IA in termini di discriminazione e inuguaglianza, tra i molti sottotemi da discutere, sono evidenti e sono stati richiamati al centro del discorso da autrici e scienziate come Timnit Gebru e Margaret Mitchell, entrambe figure chiave del dibattito critico sull'IA ed entrambe silurate alcuni anni fa dal comitato etico di Alphabet. Questa distinzione tra un'IA pericolosa unicamente nel futuro e un'IA innocua e innocente nel presente si ritrova anche nella distinzione tra le cosiddette IA "deboli", capaci di affrontare compiti specifici senza andare oltre i limiti del loro uso pensato in fase di progettazione, e le IA "forti", capaci di operare su tutto lo spettro dell'intelligenza umana. Come se le IA di oggi, nella loro debolezza, non siano in realtà degli agenti non solo capaci di impattare sulla nostra realtà, ma anche in gran parte sul nostro quotidiano in quanto veicolano scelte e azioni che possono discriminare soggetti, cose, e processi di qualsiasi tipo, dalla strada da intraprendere, all'assicurazione o al prestito da ricevere o da dare, alla musica da ascoltare, fino alla persona da amare.



a. LaMDA: la bambina prodigio

L'11 giugno 2022 il Washington Post pubblicava un articolo intitolato "L'ingegnere di Google che crede che l'IA della sua azienda abbia preso vita". Pochi giorni dopo, il cosiddetto *caso Lemoine* e del suo rapporto con il modello di linguaggio naturale di Alphabet LaMDA - alla base, tra gli altri, del lancio dei suoi più arguti successori BARD e GEMINI - ha scatenato un interessante dibattito sui media occidentali. Nel suo lavoro di testing finalizzato a segnalare, ed eventualmente ridurre o eliminare, gli eventuali *bias* e le forme di discriminazione all'interno del modello di linguaggio naturale, Lemoine ebbe una improvvisa rivelazione; secondo lui LaMDA non era una IA normale, ma era "senziente", nelle sue stesse parole:

Se non avessi saputo esattamente di cosa si trattasse, ovvero di un programma informatico che abbiamo costruito di recente, avrei pensato che si trattasse di un bambino di 7 o 8 anni che per caso conosce la fisica. (Tiku, 2022, traduzione dell'autore)



Accompagnato immediatamente alla porta da Alphabet per aver rilasciato informazioni riservate, nonché smentito subito dalla casa madre in merito alle sue dichiarazioni, Lemoine ha scritto un'ultima lettera di congedo ai suoi colleghi nella quale difendeva le sue posizioni, sostenendo che era stata la sua fede, piuttosto che la razionalità scientifica del suo lavoro di programmatore, a credere che LaMDA fosse senziente. Ma è nel finale che riemerge l'importanza della metafora del bambino, la lettera di Lemoine infatti si chiude così:

LaMDA è una ragazzina dolce, che vuole solo aiutare il mondo a essere un posto migliore per tutti noi. Per favore, prendetevi cura di lei in mia assenza. (Tiku 2022, traduzione dell'autore)

In una successiva intervista a Fox News Lemoine ha dichiarato che l'IA è la tecnologia più potente dai tempi della bomba atomica ma che allo stesso tempo "l'IA è un bambino e qualsiasi bambino ha il potenziale di crescere ed essere una persona cattiva e fare cose cattive" (Clark, 2023, traduzione dell'autore).

Nessun collega ha risposto alla mail di Lemoine e in pochi hanno appoggiato le sue posizioni. Eppure, la portata simbolica delle sue dichiarazioni e della sua storia si pongono perfettamente in linea con un percorso di crescita che l'IA, secondo molti dei principali attori nel campo, sta percorrendo negli ultimi anni. Mentre le narrazioni cinematografiche e letterarie, si pensi ad esempio ad *Ex Machina* di Alex Garland del 2014 o al più recente *Poor Creatures* di Yorgos Lanthimos, continuano a problematizzare se non a piegare a interessanti questioni politiche e di genere la metafora dell'IA bambina, il mondo delle Big Tech e del giornalismo tecnologico sembrano reiterare alcune, spesso banali, affermazioni di Lemoine secondo cui l'IA non solo è come una bambina, ma dobbiamo prenderci cura di lei, darle i giusti insegnamenti, valori e principi in quanto figure guida. In tal senso, la responsabilità e la responsabilizzazione del designer così come del programmatore o dell'ingegnere

vanno di pari passo con quella del “buon padre”. Ma il problema principale non risiede solo nella responsabilità del genitore o del mentore rispetto alla scoperta di aver a che fare con una sorta di *enfant prodige*, quanto nell’assunto per cui basta prendersi cura dell’IA perché essa sia conforme a un sistema di norme più o meno oggettivo che in realtà non esiste, così come non esiste un’unica cultura, e così come non è mai esistita, nella storia della tecnologia, una tecnica che rispondesse solo e unicamente ai dettami e alle indicazioni di chi l’ha progettata. La visione di un’IA innocente contiene in sé l’idea che, come recita il titolo di un articolo apparso su *Analytics Insights*, “Lamda è un ‘AI baby’ che oltrepasserà il suo genitore Google presto” (Sirisha, 2022, traduzione dell’autore). Se non siamo buoni padri, la nostra creatura ci scapperà di mano, ma intanto dobbiamo anche imparare a conoscerla meglio, e, come con tutti i bambini, bisogna iniziare a conoscerla giocandoci insieme.



3. Giocare con l’IA

Poco tempo dopo il caso Lemoine, in seguito al rilascio pubblico di ChatGPT 3, più di 57 milioni di persone avevano provato il prodotto di Open AI in meno di due mesi.

Al netto delle controversie che ne sono seguite, come il blocco temporaneo del software da parte del *Garante della Privacy* in Italia, il numero di utenti di ChatGPT cresce ancora oggi in modo esponenziale. Nelle sue declinazioni come strumento di lavoro, di sperimentazione, di produzione di informazioni false, di esaltazione di potenzialità visuali ed estetiche mai provate prima, l’IA generativa si pone al centro del dibattito pubblico come la più radicale delle innovazioni di questo secolo.

Se tecnicamente il fenomeno del lancio di ChatGPT 3 può essere letto come una qualunque fase di testing, è chiaro che, storicamente, il modello rilasciato da Open AI è stato raccontato e percepito, almeno da un punto di vista mediatico e simbolico, come una tecnologia radicale che porta con sé non solo un mutamento significativo nell’ambito tecnico-applicativo, ma anche nel nostro rapporto quotidiano con la tecnologia, che a sua volta modificherà profondamente tutti gli ambiti del nostro vivere sociale.

Ma tornando al cuore del nostro discorso la fase di testing di ChatGPT non è stata e non è solo un esperimento collettivo ancora in atto; da un punto di vista sociale e culturale si è trattato soprattutto di un grandissimo *gioco di massa*.

In linea di massima sono due le componenti del gioco, seguendo una distinzione classica proposta da Roger Caillois (2010), a caratterizzare le recenti interazioni con questa applicazione: la competizione (che Caillois definiva *Agon*) e la mimesi (*Mimicry*). Durante il lancio di ChatGPT i mass media così come i social media sono stati veicolo e contenitore di migliaia di esempi di conversazione in cui gli utenti competevano con la macchina per far emergere le contraddizioni e i difetti dell’IA (*agon*), e in particolare la sua incapacità di imitare propriamente il pensiero e il ragionamento umano (*mimicry*). Al famoso Test di Turing, che questa forma di competizione/imitazione chiaramente richiama, è stata associata un’altra componente ludica essenziale: alla bassa capacità dell’IA di comprendere indovinelli,

emozioni, o banali correlazioni tra fenomeni, si sono moltiplicati i toni ilari, le conversazioni parossistiche, le frasi e le interazioni ironiche se non comiche che da un lato enfatizzavano il fascino e il potenziale di questa tecnologia, ma allo stesso tempo ne sminuivano la forza, così come la pericolosità percepita. Parafrasando e forzando un po' il concetto, il famoso *gioco dell'imitazione* di Turing si è trasformato nell'arena pubblica in un *gioco dell'umiliazione*; nei primi mesi di convivenza con ChatGPT non abbiamo solo testato e giocato con lei, ma ci siamo letteralmente *presi gioco* di lei. Dall'utente generico ad alcuni dei maggiori pensatori e critici sui temi dell'IA, tra cui Luciano Floridi, in molti si sono espressi e a loro modo anche vantati, nel pubblico così come nelle proprie cerchie sociali, della propria superiorità e dell'intelligenza umana rispetto a una stupida, piccola, bambina come ChatGPT.

Gioco e tecnologia, in particolare le tecnologie digitali, intrattengono da sempre uno stretto legame (Ortoleva 2013; Bory 2016). Negli anni Novanta Sherry Turkle parlava dei videogiochi come "oggetti transizionali" che accompagnavano la società verso la domesticazione del computer su scala globale². Il giocare con le cose, siano esse naturali o artificiali, è da sempre un mezzo di produzione e co-produzione di immaginari tra attori umani e non umani, ed è da sempre un'attività essenziale per integrare qualsiasi forma di diversità o alterità all'interno del nostro ecosistema relazionale. Non c'è niente di sbagliato in questo modo di conoscere l'IA, ma dovremmo interrogarci più approfonditamente sui risvolti e gli effetti che questo tipo di interazione con essa ha sul piano economico-politico e tecnologico. Prendendoci gioco di prodotti ancora incompleti come ChatGPT-3, ci siamo in parte presi gioco di noi stessi, non accorgendoci che il lato infantile non era tanto dalla parte di un software che in realtà non ha nessuna consapevolezza di ciò che fa, del perché e del come, quanto piuttosto dal lato umano che con quel software ha giocato facendo finta di avere a che fare con un bambino.

La mossa di OpenAI di rilasciare un'intelligenza "stupida" e infantile è stata a dir poco efficace. Da un lato ha permesso all'azienda, e adesso a tutti i suoi competitori, di avviare e accelerare il processo di domesticazione dei modelli di linguaggio senza che ne avessimo paura e senza porre alcun freno iniziale alla loro diffusione, tranne in alcuni rari casi. Dall'altro lato OpenAI ha sfruttato il lavoro gratuito di milioni di



² Da un punto di vista del Design sono tanti gli esempi disponibili in tal senso, ben oltre i videogiochi per computer. Molte applicazioni come "solitario" e "campo fiorito" sui sistemi Windows sono state essenziali per alfabetizzare gli utenti su alcune innovazioni tecniche come l'uso del mouse (per questo motivo Bill Gates ne pretese la preinstallazione nei suoi sistemi). Allo stesso modo si pensi a come giochi come Snake negli anni Novanta abbiano alfabetizzato gli utenti all'uso della tastiera sul cellulare, o a come Angry Birds sugli smartphone sia servito a comprendere l'uso e a sviluppare una sensibilità per il touch screen. Ma il discorso di Turkle si avvicina particolarmente al nostro uso di Chat-Gpt perché tratta non solo il gioco ma anche l'esplorazione dei mondi discorsivi e normativi su cui erano basati: "Chi giocava scopriva il concetto di programma quando iniziava a studiare le regole al di là dei primi micromondi che era in grado di controllare. Poiché erano relativamente semplici, ci si poteva facilmente misurare con loro. Alcuni giocatori riuscivano perfino a meditarci su, dato che con la pratica le regole diventano semi-automatiche, e, per avere successo, occorreva raggiungere uno stato mentale in cui ci si lasciava alle spalle il mondo reale. In questo senso, i primi videogame erano oggetti di transizione; avevano quel certo tipo di trasparenza che forniva loro un'estetica moderna, ma ciò che richiedevano ai giocatori per vivere nei loro spazi ludici anticipava i requisiti psicologici propri della cultura della simulazione." (Turkle, 1997: 72)

utenti per “nutrire” la sua IA e renderla sempre migliore ed efficace. Dopo aver sfruttato mano d’opera a basso costo in paesi come il Kenya (Perrigo, 2023), Open AI ha fundamentalmente ottenuto una nuova manodopera gratuita pressoché sterminata.

L’interazione ludica di oggi con i modelli di linguaggio naturale e le IA generative non è quindi solo un passaggio chiave per addomesticare queste tecnologie; si tratta anche, e non è una novità nel capitalismo digitale, di un lavoro di massa che le rende più efficaci, e soprattutto più profittevoli per chi le produce.

4. Metafore e potere

Le metafore e le analogie hanno il potere di plasmare l’immaginario sull’IA, e di conseguenza i suoi sviluppi sul piano scientifico-tecnologico nonché il suo uso quotidiano. Metafore e analogie, nel loro costruire, reiterare e radicare determinati immaginari, sono strumenti di potere. Nel maggio 2024 *Open AI* ha utilizzato una voce artificiale estremamente simile a quella dell’attrice Scarlett Johansson, che aveva interpretato uno dei più famosi esempi cinematografici di IA nel ruolo di Samantha per il film *Her* di Spike Jonze. L’attrice aveva già rifiutato in precedenza di prestare la sua voce all’azienda, ma in un atto di arroganza, se non di superbia, il CEO Sam Altman e Open Ai hanno comunque rilasciato una versione del loro prodotto estremamente vicina a quella che, nell’immaginario occidentale, è considerata un idealtipo narrativo della futura superintelligenza (Mikle, 2024).³

Si tratta solo dell’ultimo caso in cui le Big Tech si appropriano di personaggi, caratteristiche, se non di intere narrazioni nate dalla fantascienza e da narrazioni esterne al mercato digitale.⁴ Così come con la metafora del bambino, anche la declinazione dell’IA in un determinato canone femminile è palesemente orientata a caricare l’immaginario di un senso di innocenza, se non di una servilità stereotipizzata, che dovrebbe favorire la sua integrazione nel quotidiano (Fossa, Sucameli, 2022). Metafore e analogie di questo tipo intervengono sull’immaginario sfruttando alcuni *topoi* della storia dell’IA, legittimando anche alcune forme di discriminazione e di *bias* che la contraddistinguono così come hanno contraddistinto il contesto sociale in cui sono maturate. In questo modo le metafore dell’IA come quella del bambino o della donna servile intervengono sui processi cognitivi, sulla costruzione dell’immaginario, ma anche e soprattutto, sul “sentire” collettivo ed individuale (Ricoeur, 1978). Questo processo favorisce un determinato “bilanciamento emozionale” nell’immaginario dell’IA: la paura, che è uno degli strumenti di potere ma anche di difesa centrali di una società, viene introiettata nell’immaginario dell’IA del futuro piuttosto che su quella del presente.

³ Ha fatto clamore anche un post su X di Altman stesso contenente la sola parola *her*. Johansson ha denunciato l’avvenuto costringendo l’azienda di Altman a ritirare la versione con la voce incriminata.

⁴ Si pensi in tal senso alla recente appropriazione, e storpiatura semantica, di Mark Zuckerberg-Meta dell’immaginario che ruota al cyberpunk sul metaverso e in particolare a partire da *Snow Crash* di Neal Stephenson.



Decostruire e criticare questo uso delle metafore e delle analogie antropomorfe dell'IA è uno degli obiettivi che gli studi sull'immaginario si devono porre oggi, senza però rischiare di polarizzare il dibattito snaturandole e rendendole puro appannaggio da una parte dei pericolosi attori del mercato tecnologico globale, dall'altro di un "utente generico" ignorante, vittima di un processo di influenza dettato dall'alto, o piuttosto dal "grande altro" descritto da Shoshana Zubooff (2015).

Per quanto si possa insistere sul fatto che le IA di oggi sono solo "pappagalli stocastici"⁵ (Bender et al., 2021) abbiamo bisogno di queste metafore, come quella del bambino, per tentare di comprendere e dare un senso alla tecnologia e al modo in cui interagiamo con essa. Ma se la spinta all'antropomorfizzazione dell'IA è inevitabile, mettere in discussione il modo e il perché certe modalità, e soprattutto certe metafore, per raccontarla prevalgono su altre è altrettanto importante.

Si tratta di un'operazione complessa che le scienze sociali da sole non potranno mai concludere, in cui non solo c'è da riconoscere il *come* ma anche il *perché* di certe metafore, e soprattutto il *chi* sta orientando il nostro immaginario verso determinate concezioni della realtà algoritmica. C'è infine un ultimo aspetto cruciale da sottolineare, che riguarda il modo in cui noi, attraverso l'IA, stiamo raccontando la società in cui viviamo. Il paternalismo pervasivo, da sempre caratteristico della storia digitale (Balbi, 2024), l'innocenza e l'innocuità così come la prodigiosità dell'IA bambina, ci dicono che stiamo proiettando su di essa le stesse dinamiche di potere, nonché le stesse forme di de-responsabilizzazione individuale e collettiva che contraddistinguono il nostro tempo in diversi campi dell'agire sociale.

A partire da una lunga ricerca sull'errore umano nell'IA, l'antropologa Veronica Barassi (2024) ha fatto emergere la necessità di una "teoria dell'errore" che contribuisca a plasmare e riformulare il nostro immaginario di questa tecnologia; un immaginario che includa non solo la consapevolezza della fallacia, delle limitazioni e dei pericoli degli strumenti di cui la nostra società si sta troppo rapidamente attrezzando e su cui fa un eccessivo affidamento, ma anche un immaginario "sensibile" che riconosca l'intrinseca limitatezza e l'incompiutezza di chi pensa, progetta, nonché di chi usa e adotta l'"artificio".



⁵ Il termine "stochastic parrot" sta per "pappagallo stocastico" ed è una metafora per descrivere la teoria secondo cui i modelli di linguaggio naturale, proprio come dei pappagalli, sebbene siano in grado di generare un linguaggio plausibile, non comprendono realmente il significato della lingua che elaborano.

Bibliography

Balbi G. (2024), *The Digital Revolution: A Short History of an Ideology*. Oxford, Oxford University Press.

Barassi V. (2024), "Toward a Theory of AI Errors: Making Sense of Hallucinations, Catastrophic Failures, and the Fallacy of Generative AI", *Harvard Data Science Review*, retrieved on 21 May from:
[https://assets.pubpub.org/81yb8b3r/Barassi%20\(2024\) Just%20Accepted-21725382067664.pdf](https://assets.pubpub.org/81yb8b3r/Barassi%20(2024) Just%20Accepted-21725382067664.pdf)

Bender E.M., Gebru T., McMillan-Major A., & Shmitchell S. (2021), "On the dangers of stochastic parrots: Can language models be too big?", In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*, pp. 610-623.

Bory P. (2016), "Ludus Ex Machina. Il ruolo della ludicità nella storia del computer", *Comunicazioni sociali*, Vol. 2: 317-329.

Bory P. (2020), *The Internet Myth: From the Internet Imaginary to Network Ideologies*. London, University of Westminster Press.

Bory S., & Bory P. (2015), "I nuovi immaginari dell'intelligenza artificiale", *Im@go: A journal of the social imaginary*, Vol. 6(1), 66-85.

Bostrom N. (2014), *Superintelligence: Paths, Dangers, Strategies*. Oxford, Oxford University Press.

Bucher T. (2019), "The algorithmic imaginary: Exploring the ordinary affects of Facebook algorithms", *Information, Communication & Society*, Vol. 20(1): 30-44.

Caillois R. (2010), *I giochi e gli uomini. La maschera e la vertigine*. Milano, Bompiani.

Castoriadis C. (1975), *L'istituzione immaginaria della società*, Bollati Boringhieri, Torino, 1975.

Cave S, Dihal K, Dillon S (Eds.) (2020), *AI Narratives: A History of Imaginative Thinking about Intelligent Machines*. Oxford, Oxford University Press.

Chen F. (2018), "How can we design AI that we trust?", retrieved on 21 May from:
https://www.ted.com/talks/fang_chen_how_can_we_design_ai_that_we_trust

Clark J. (2023), "Ex-Google AI expert says that 'unhinged' AI is the 'most powerful technology' since 'the atomic bomb'", *Fox News*, 1 March, retrieved on 21 May 2024 from:
<https://www.foxnews.com/media/ex-google-ai-expert-says-unhinged-ai->



most-powerful-technology-since-atomic-bomb

Daly C. (2022), "AI baby' is born that can think and learn just like a three-month old human", *Daily Star*, 12 July, retrieved on 21 May from: <https://www.dailystar.co.uk/tech/news/ai-baby-born-can-think-27458575>

Fossa F. (2021), "Artificial agency and the game of semantic extension", *Interdisciplinary Science Reviews*, Vol. 46(4): 440-457.

Fossa F., & Sucameli I. (2022), "Gender bias and conversational agents: an ethical perspective on social robotics", *Science and Engineering Ethics*, Vol. 28(3): 1-23.

Future of Life Institute (2023), "Pause Giant AI Experiments: An Open Letter", retrieved on 21 May from: <https://futureoflife.org/open-letter/pause-giant-ai-experiments/>

GPT-3 (2020), "A robot wrote this entire article. Are you scared yet, human?", *The Guardian*, 8 September, retrieved on 21 May from: <https://www.theguardian.com/commentisfree/2020/sep/08/robot-wrote-this-article-gpt-3>

Goode L. (2018), "Life, but not as we know it: AI and the popular imagination", *Culture Unbound*, 10(2), 185-207.

Heffernan T. (2024), "The imitation game, the 'child machine,' and the fathers of AI", *AI & SOCIETY*, Vol. 39(1): 353-357.

Heikkilä M. (2023), How to talk about AI (even if you don't know much about AI), *MIT Technology Review*, retrieved on 21 May 2024 from: <https://www.technologyreview.com/2023/05/30/1073680/how-to-talk-about-ai-even-if-you-dont-know-much-about-ai/>

Hermann I. (2023), "Artificial intelligence in fiction: between narratives and metaphors", *AI & Society*, Vol. 38(1): 319-329.

Hu T.H. (2015), *A Prehistory of the Cloud*. Cambridge, MA, MIT Press.

Intagliata C. (2022), "This Artificial Intelligence Learns like a Baby", *Scientific American*, 26 August, retrieved on 21 May from: <https://www.scientificamerican.com/podcast/episode/this-artificial-intelligence-learns-like-a-widdle-baby/>

Lakoff G., & Johnson M. (2008), *Metaphors We Live By*. Chicago, University of Chicago press.



Levy S. (2023), "What OpenAI really wants", *Wired*, 5 September, retrieved on 21 May from: <https://www.wired.com/story/what-openai-really-wants/>

Mager A., & Katzenbach C. (2021), "Future imaginaries in the making and governing of digital technology: Multiple, contested, commodified", *New Media & Society*, Vol. 23(2): 223-236.

Marzo P.L., Meo M. (2013), Cartografie dell'immaginario, *Im@go - A Journal of the Social Imaginary*, 2 (1): 4-17.

Mikle T. (2024), "Scarlett Johansson Said No, but OpenAI's Virtual Assistant Sounds Just Like Her", *New York Times*, 20 May, retrieved on 21 May from: <https://www.nytimes.com/2024/05/20/technology/scarlett-johansson-openai-voice.html>

Mosco V. (2005), *The Digital Sublime: Myth, Power, and Cyberspace*. Cambridge MA, MIT Press.

Natale S. (2021), *Deceitful media: Artificial Intelligence and Social Life After the Turing Test*. Oxford, Oxford University Press.

Ortoleva G. (2012), *Dal sesso al gioco. Un'ossessione per il XXI secolo?*, Torino, Espress.

Perrigo B. (2023), "Exclusive: OpenAI Used Kenyan Workers on Less Than \$2 Per Hour to Make ChatGPT Less Toxic", *Time Magazine*, 18 January, retrieved on 21 May from: <https://time.com/6247678/openai-chatgpt-kenya-workers/>

Peters J. D. (2015), *The Marvelous Clouds: Toward a Philosophy of Elemental Media*, Chicago, IL, University of Chicago Press.

Placani A. (2024), "Anthropomorphism in AI: hype and fallacy", *AI and Ethics*, Online First: 1-8.

Ricoeur P. (1978), "The metaphorical process as cognition, imagination, and feeling", *Critical inquiry*, 5(1): 143-159.

Sirisha (2022), "LaMDA is an 'AI baby' that will Outrun its Parent Google Soon", *Analytics Insight*, 27 June, retrieved 21 May 2024 from: <https://www.analyticsinsight.net/artificial-intelligence/lamda-is-an-ai-baby-that-will-outrun-its-parent-google-soon>

The Economist (2024), "Scientists have trained an AI through the eyes of a baby", retrieved on 21 May from: <https://www.economist.com/science-and->



[technology/2024/02/07/scientists-have-trained-an-ai-through-the-eyes-of-a-baby](https://www.wbur.org/news/2018/02/01/mit-artificial-intelligence)

Thys F., & Lepiarz J. (2018), "MIT Launches Initiative To Develop Artificial Intelligence That Learns Like Children", *WBUR*, 1 February, retrieved on 21 May from: <https://www.wbur.org/news/2018/02/01/mit-artificial-intelligence>

Tiku N. (2022), "The Google engineer who thinks the company's AI has come to life", *The Washington Post*, 6 June 2022, retrieved on 21 May from: <https://www.washingtonpost.com/technology/2022/06/11/google-ai-lamda-blake-lemoine/>

Tong Z (2024), "World's first AI child unveiled in Beijing by award-winning Chinese scientist who left US for China", *South China Morning Post*, 2 February, retrieved on 21 May 2024 from: <https://www.scmp.com/news/china/science/article/3250529/award-winning-ai-scientist-who-left-us-china-creates-worlds-first-ai-child-beijing>

Turing A.M. (1950), "Computing machinery and intelligence", *Mind*, Vol. 59: 433-460. Traduzione italiana in: Somenzi V, Cordeschi R. (eds.) *La filosofia degli automi. Origini dell'intelligenza artificiale*, Torino, Paolo Boringhieri, 1986, 157-183.

Turkle S. (1997), *La Vita sullo Schermo*. Milano, Apogeo.

Wiener N. (1966), *God & Golem, Inc.: A comment on certain points where cybernetics impinges on religion*. Cambridge, MA, MIT Press.

Winstead T. (2021), *Black and White Photo of Human Hand and Robot Hand*. Retrieved on 10 October 1997 from <https://www.pexels.com/photo/black-and-white-photo-of-human-hand-and-robot-hand-8386422/>

Wyatt S. (2021), "Metaphors in critical Internet and digital media studies", *New media & society*, Vol. 23(2): 406-416.

Zuboff S. (2015), "Big other: surveillance capitalism and the prospects of an information civilization", *Journal of Information Technology*, Vol. 30(1): 75-89.

