

## **ChatGPT come rischio o opportunità? Le lenti dell'Information Literacy per comprendere e usare criticamente l'Intelligenza Artificiale Generativa. Un report dell'Unesco**

Laura Testoni

- *Intelligenza Artificiale Generativa come nuovo synthetic media.*

L'intelligenza artificiale generativa, in quanto strumento in grado di produrre contenuti, sta diventando un nuovo media sintetico di massa in grado di rappresentare una nuova fonte epistemica.

Per questa ragione interroga in modo sostanziale i professionisti dell'informazione e le biblioteche, in particolare quelle accademiche e di ricerca.

Con l'espressione "intelligenza artificiale (IA) generativa e conversazionale" intendiamo quegli strumenti e tool in grado di produrre contenuti a partire da un prompt immesso dall'utente.

Il *prompt* è una frase redatta in linguaggio naturale da un utilizzatore, che viene fornita al sistema per indirizzarlo verso la generazione di contenuti. Output del prompt sono testi compiuti, grafici o immagini.

Premesso che la tecnologia alla base dell'IA generativa è strutturata da molti fattori non semplici da spiegare e comprendere<sup>1</sup>, occorre sapere che alla base di questi strumenti ci sono complesse architetture di *Machine learning*<sup>2</sup> che "apprendono" attraverso la suddivisione in frammenti (cd. "token") di immensi corpora di testi e la loro successiva elaborazione in termini di occorrenze statistiche.

È in questo senso che i modelli basati sull'elaborazione di enormi quantità di contenuti testuali sono detti "Large Language Models" (LLMs)<sup>3</sup>.

I processi di elaborazione dei contenuti generati dagli strumenti di IA consistono nell'apprendere le relazioni statistiche tra frammenti di testo attraverso algoritmi detti Transformer<sup>4</sup>.

---

<sup>1</sup> "Non so come funzionino veramente ChatGPT e i suoi molti cugini, non lo sa ancora nessuno". Nello Cristianini (2024), *Machina Sapiens : l'algoritmo che ci ha rubato il segreto della conoscenza*, Bologna : Il Mulino, p. 7.

<sup>2</sup> "...Quando si produce software i programmatori generano un codice che include un set di istruzioni sequenziali e regole che producono un risultato dato. Il *Machine learning* opera in modo differente, cioè "apprende" attraverso una procedura di training, intercettando i patterns presenti su enormi quantità di dati somministrati al sistema. Le procedure di Machine learning costruiscono da sole il proprio set di istruzioni per ottenere l'output". Cfr. M. Kurpicz-Briki (2023). *An Introduction to Machine Learning*. In: *More than a Chatbot*. Springer, Cham. doi 10.1007/978-3-031-37690-0\_2. Sintesi e traduzione mia.

<sup>3</sup> Il termine *Large Language Model* (LLM) si riferisce generalmente a un modello di linguaggio che ha un ampio numero di parametri ed è allenato su dataset massivi. [...] in termini di IA i parametri sono gli aspetti del modello appresi dai dati su cui il sistema si allena. Più ampio è il numero di parametri più il modello può apprendere. I LLM moderni hanno miliardi di parametri. Ad esempio, si stima che ChatGPT3 abbia 175 miliardi di parametri. Cfr T. Amaratunga (2023). *What Makes LLMs Large?*. In: *Understanding Large Language Models: learning their underlying concepts and technologies*, doi: 10.1007/979-8-8688-0017-7.

<sup>4</sup> Transformer è un'architettura, un modello che rivoluziona la creazione di linguaggio naturale con un meccanismo "attention-based" che consente di "pesare" l'importanza di differenti parole o lemmi catturandone le interdipendenze e interrelazioni a lungo raggio. Ciò permette di comprendere il contesto

I sistemi di IA generativa producono contenuti in modo predittivo, prevedendo cioè statisticamente le associazioni tra frammenti di testo, a partire anche dal loro contesto<sup>5</sup> e da milioni di altri parametri.

L'infrastruttura tecnologica alla base delle applicazioni di intelligenza artificiale consiste in processori e super computer aggregati in enormi data centre sparsi per il mondo.

Tra tutti i sistemi di IA generativa conversazionale, il più noto è ChatGPT lanciato dall'Azienda Open AI<sup>6</sup> il 22 ottobre 2022 e ad oggi [giugno 2024] arrivato, nella versione gratuita e aperta a tutti, alla versione "o" (come "Omnia").

I sistemi di IA generativa sono attualmente impiegati in moltissimi ambiti<sup>7</sup>: medicina, business e finanza, servizi legali, scrittura creativa di ogni tipo, training ed educazione (es. preparazione di test o tutoring online), creazione o debugging di codice in ambito informatico, servizi di vendita e marketing, servizi bancari, ma anche nella vita quotidiana<sup>8</sup>. Alla base dell'apprendimento dei sistemi di IA c'è l'elaborazione di corpora di testi molto vasti<sup>9</sup>

- *IA in ambito accademico e nella produzione scientifica.*

Anche in ambito accademico è documentato, come vedremo, l'utilizzo di sistemi di intelligenza artificiale, soprattutto in contesti quali: l'analisi dei set di dati, la review di corpus di articoli scientifici, la redazione di riassunti e sintesi, la generazione di bibliografie.

---

della conversazione in corso e produrre risposte coerenti, in linea con il dialogo precedente, generando una risposta contestualmente appropriata. Cfr. A. Kulkarni, A. Shivananda, A. Kulkarni, D. Gudivada (2023). The ChatGPT Architecture: An In-Depth Exploration of OpenAI's Conversational Language Model. In: Applied Generative AI for Beginners. Apress, Berkeley, CA. doi 10.1007/978-1-4842-9994-4\_4.

<sup>5</sup> Segnaliamo alcuni testi sul funzionamento dell' IA generativa conversazionale: Nello Cristianini (2024) *Machina Sapiens : l'algoritmo che ci ha rubato il segreto della conoscenza*, cit.; Gino Roncaglia (2023) *Intelligenze artificiali generative e mediazione informativa. Una introduzione. Biblioteche oggi Trends*, vol. 9, n. 1 (giugno 2023) p. 13-26 doi: 10.3302/2421-3810-202301-013-1; Gino Roncaglia (2023) *L'architetto e l'oracolo. Forme Digitali Del Sapere Da Wikipedia a ChatGPT*. Roma Bari : Laterza; Luciano Floridi, (2023), *AI as Agency Without Intelligence: On ChatGPT, Large Language Models, and Other Generative Models* (February 14, 2023). *Philosophy and Technology*, 2023 <<https://ssrn.com/abstract=4358789>>.

<sup>6</sup> OpenAI è al terzo posto nella classifica delle start up di maggior valore al mondo dopo ByteDance (TikTok) e Space X; ha ricevuto nel 2023 dieci miliardi di dollari da Microsoft che è il suo maggior finanziatore. Cfr. Sara Bichicchi, *Intelligenza artificiale, OpenAI: 1,6 miliardi di dollari di fatturato nel 2023. E ora potrebbe valere 100 miliardi*, *Milano finanza* 2/1/2024 <<https://www.milanofinanza.it/news/openai-nel-2023-fatturato-a-1-6-miliardi-di-dollari-e-ora-la-startup-di-sam-altman-potrebbe-valere-100-202401011646045378>>.

<sup>7</sup> Per un elenco molto esaustivo e ricco di dettagli si cita: ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. Partha Pratim Ray, *Internet of Things and Cyber-Physical Systems* 3 (2023) 121–154 doi 10.1016/j.iotcps.2023.04.003, specialmente sez.6 6. Applications across various domains.

<sup>8</sup> Francesca Paris, Larry Buchanan, 35 Ways Real People Are Using A.I. Right Now, *The New York Times*, April 14, 2023.

<sup>9</sup> "ChatGPT è addestrato utilizzando circa 500 miliardi di token, la maggioranza di questi provenienti da pagine web, il resto da libri disponibili online e da Wikipedia" Cfr. Nello Cristianini, *Machina Sapiens : l'algoritmo che ci ha rubato il segreto della conoscenza*, cit. p. 136.

Sono note le prese di posizione e i *caveat* da parte delle maggiori riviste scientifiche come *Nature*<sup>10</sup> e *Science*<sup>11</sup> che hanno ribadito da una parte la questione dell'Authorship e il tema della *scientific integrity* e, dall'altra, la necessità di approfondire il ruolo dell'intelligenza artificiale nella produzione scientifica.

Cope (*Committee on Publication Ethics*), l'organizzazione no-profit che supporta autori ed editori sui temi dell'etica nella comunicazione scientifica, ha pubblicato nel 2023 un suo *Position statement*<sup>12</sup> in cui sottolinea che "gli strumenti di IA non possono aderire ai requisiti per l'Authorship in quanto non assumono responsabilità per i lavori sottoposti e, come entità non-giuridiche, non possono asserire la presenza o assenza di conflitti di interesse né gestire copyright o accordi di licenza"<sup>13</sup>.

Il dibattito resta comunque aperto, e alcuni editori scientifici stanno rivedendo le proprie policy editoriali<sup>14</sup>.

In ambito strettamente bibliotecario, il tema dell'Authorship nel contesto dell'IA ha implicazioni bibliografiche<sup>15</sup>; inoltre i maggiori database citazionali<sup>16</sup> (e non solo<sup>17</sup>) hanno già o stanno introducendo strumenti di IA basati su Large Language Models incorporati e complementari alla ricerca tradizionale o booleana, capaci di fornire, a un quesito espresso in linguaggio naturale, una risposta accurata corredata dagli stessi dati bibliografici (o fattuali) indicizzati nella risorsa.

In questo quadro emergono alcuni aspetti da sottolineare: come noto gli strumenti di IA che si allenano sui dataset massivi necessari sulla base dei Large Language Models, non sono immuni da bias, errori, stereotipi ed etnocentrismi, ereditati esattamente dai medesimi corpora testuali su cui è basato il loro apprendimento<sup>18</sup>.

---

<sup>10</sup> Thorp H. Holden (2023). ChatGPT is fun, but not an author. *Science* 379 Issue 6630 (2023), p. 313 <<https://doi.org/10.1126/science.adg7879>>. Rispetto agli errori di ChatGPT nella compilazione di bibliografie: W.H. Walters, E.I. Wilder (2023) Fabrication and errors in the bibliographic citations generated by ChatGPT. *Nature Scientific Reports* 13, 14045. doi: 10.1038/s41598-023-41032-5.

<sup>11</sup> Eva A.M. Van Dis Eva, Johan Bollen, Willen Zuidema, Robert van Rooij, Claudi Bockting (2023) ChatGPT: five priorities for research, *Nature* 614, 224-226 doi: 10.1038/d41586-023-00288-7.

<sup>12</sup> COPE Authorship and AI Tools. COPE position statement <<https://publicationethics.org/cope-position-statements/ai-author>>.

<sup>13</sup> Cfr. Ibidem

<sup>14</sup> Chris Stokel-Walker (2023), ChatGPT listed as author on research papers: many scientists disapprove. At least four articles credit the AI tool as a co-author, as publishers scramble to regulate its use, *Nature* 613, 620-621 doi: 10.1038/d41586-023-00107-z.

<sup>15</sup> Cfr. Maurizio Lana. The Agency of Artificial Intelligence Systems. *DigitCult - Scientific Journal on Digital Cultures*, [S.l.], v. 7, n. 1, p. 67-78, July 2022. ISSN 2531-5994. doi: 10.36158/97888929552576.

<sup>16</sup> Scopus dell'editore Elsevier ha posto in commercio quest'anno ScopusAI; Web of science di Calarivate Analytics ha annunciato il lancio di un assistente IA nel database per settembre 2024.

<sup>17</sup> Il database Statista, che riaggrega ed espone dati da numerosi fonti statistiche globali consente un'interrogazione in linguaggio naturale offrendo una risposta articolata e corredata da suoi propri dati. Tale modalità è offerta gratuitamente per le prime due settimane successive a lancio del servizio cfr. <<https://www.statista.com/research-ai/>>. Occorre osservare che i database accademici offrono il servizio "IA" ovvero la possibilità di essere interrogati e fornire output in linguaggio naturale quasi sempre come servizi aggiuntivi a pagamento.

<sup>18</sup> Su questo aspetto, e più in generale sui limiti dell' IA generativa conversazionale, la letteratura è ormai sterminata. Si veda, su tutti: Bender Emily M., Timnit Gebru, McMillan-Major Anelina, Shmitchell

Le istituzioni della conoscenza, tra cui le biblioteche, che sono orientate su una cultura dell'information literacy e quindi praticano un approccio critico verso le fonti informative e medial, non possono ignorare, e non ignorano<sup>19</sup> l'impatto dirompente dell' IA generativa nel panorama documentale.

- *Media Information Literacy e Intelligenza Artificiale. Un documento Unesco*

È appena stato pubblicato, in questo contesto, un documento Unesco<sup>20</sup> che presenta molti spunti di riflessione sul rapporto tra Information Literacy e ChatGPT (e più in generale strumenti di IA generativa). Lo illustriamo nelle righe che seguono.

Il documento ha un titolo programmatico: "*User empowerment through media and information literacy responses to the evolution of generative artificial intelligence*"; pubblicato nel 2024 declina in modo sistematico le risposte che la Media Information Literacy (MIL)<sup>21</sup> può offrire rispetto al dilagare dei contenuti prodotti o co-redatti da IA generativa. In questa chiave Unesco propone di utilizzare le lenti dell'Information literacy per gestire il cambiamento introdotto dai sistemi di IA nel contesto informativo.

Più precisamente le strategie e le policy della MIL sono ritenute la chiave per fornire alle persone gli "attrezzi" necessari per mitigare i rischi e valorizzare le opportunità che l'IA porta con sé. Promuovere la MIL risponde allora alla necessità di consentire alle persone di partecipare in modo informato alle conversazioni in corso sugli usi dell'IA nelle loro vite, influenzando così gli sviluppi di questa tecnologia<sup>22</sup>.

La prima mossa è senz'altro non sottovalutare queste tecnologie, pervasive perché relativamente gratuite (chiunque previa registrazione può accedere a ChatGPT), semplici da usare e francamente sorprendenti nell'efficacia e naturalezza degli output linguistici; occorre inoltre considerare che la MIL, con il suo approccio olistico e ricomprensivo differenti literacy è lo strumento adatto per affrontare questa sfida.

---

Shmargaret (2021), On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?. In Conference on Fairness, Accountability, and Transparency (FAccT '21), March 3–10, 2021, Virtual Event, Canada. ACM, New York, NY, USA, <<https://doi.org/10.1145/3442188.3445922>>;

<sup>19</sup> Cfr IFLA FAIFE Committee on Freedom of Access to Information and Freedom of Expression (2020) IFLA Statement on Libraries and Artificial Intelligence 2020 <https://repository.ifla.org/handle/123456789/1646>. Segnaliamo il commento di Maurizio Lana sul documento IFLA Faife in: Lana, Maurizio (2024), Biblioteche e intelligenza artificiale: questioni e ipotesi di lavoro, in: Maurizio Marco, Vivarelli Fabrizio, Mellia Lamberti, Biblioteche, lettura, intelligenza artificiale, Milano : Bibliografica, 2024. Segnaliamo che ACRL, l'associazione americana delle biblioteche di college e accademiche ha posto l'IA e la IA literacy come prima "trend topic" bibliotecaria del 2024: Research Planning and Review Committee, A. (2024). 2024 Top Trends in Academic Libraries: A Review of the Trends and Issues. *College & Research Libraries News*, 85(6), 231. doi: 10.5860/crln.85.6.231.

<sup>20</sup> UNESCO (2024) Frau-Meigs, Divina, User empowerment through media and information literacy responses to the evolution of generative artificial intelligence (GAI) <<https://unesdoc.unesco.org/ark:/48223/pf0000388547>>

<sup>21</sup> Come noto, in ambito Unesco (e IFLA) la Media information literacy è un termine ombrello che contiene e compendia diverse competenze, tra cui la Information literacy cfr. UNESCO (2013) Global Media and Information Literacy Assessment Framework: Country Readiness and Competencies prepared by UNESCO Communication and Information Sector <<https://unesdoc.unesco.org/ark:/48223/pf0000224655>>

<sup>22</sup> Cfr UNESCO (2024), cit, passim, p 3. Naturalmente va considerato che gli sviluppi delle tecnologie di IA sono determinati da fattori scientifici, economici e giuridici globali.

Occorre -sottolinea il report- essere consapevoli delle veloci accelerazioni nel mercato del lavoro<sup>23</sup> che chiamano in causa l'IA e le interazioni uomo-macchina (chatbox, assistenti IA): in questo senso un approccio critico e la familiarità con i sistemi di IA favorisce l'occupabilità e deve diventare parte integrante dell'aggiornamento permanente<sup>24</sup>.

Il report suggerisce alcune risorse che possono essere di supporto ad un approccio critico verso gli strumenti di IA che generano testi e notizie, come ad esempio GPTZero <<https://gptzero.me/>> un tool che dichiara di intercettare se un testo è generato da una IA oppure no; Vera.ai <<https://www.veraii.eu/home>>, parte del progetto *VERification Assisted by Artificial Intelligence* promosso dall'UE nell'ambito del progetto Horizon Europe che studia soluzioni e tecniche contro la disinformazione avanzata realizzando report e prototipi, come ad esempio InVid un fake news debunker utilizzabile in forma di estensione per Chrome.

All'interno del report viene proposto un quadro concettuale per la MIL che sia compatibile con le competenze necessarie per usare criticamente e consapevolmente gli strumenti di IA generativa. Il quadro concettuale<sup>25</sup> di articola in quattro punti:

- Conoscenza

Conoscere le potenzialità e i rischi dell'IA, gli attori geopolitici implicati, i benefici e i disordini informativi che essa comporta, monitorarne l'impronta ecologica<sup>26</sup>.

- Capacità

Usare responsabilmente gli strumenti di AI, essere in grado di fare un discrimine tra mass media, social media e synthetic media (comprendendo anche le loro interazioni), interagire con agenti artificiali padroneggiando la tecnica necessaria per porre domande attraverso un prompt<sup>27</sup>.

- Attitudini

Evitare i rischi sistemici (manipolazione, viralità, automaticità, privacy, furto di dati) connessi con i sistemi di IA conversazionale, identificando e riportando contenuti dannosi (hate speech, bias etc.).

- Valori

---

<sup>23</sup> Un report dell' IMF sottolinea come la crescente adozione dell'IA ed esposizione ad essa può generare cambiamenti nel mercato del lavoro con grandi differenze tra le nazioni: dove l'IA offre vantaggi in termini di produttività può anche porre rischi sulle riallocazioni in ambito lavorativo e ampliare le diseguaglianze già in essere. Cfr Cazzaniga and others. 2024. "Gen-AI: Artificial Intelligence and the Future of Work." IMF Staff Discussion Note SDN2024/001, International Monetary Fund, Washington, DC. <<https://www.imf.org/en/Publications/Staff-Discussion-Notes/Issues/2024/01/14/Gen-AI-Artificial-Intelligence-and-the-Future-of-Work-542379>>.

<sup>24</sup> Cfr. UNESCO (2024), cit. pp. 6 e 8 e segg.

<sup>25</sup> Cfr ibidem p. 13.

<sup>26</sup> Come noto, i grandi datacenter e le infrastrutture di calcolo che permettono all'IA Generativa di funzionare sono enormemente energivori. È stato valutato che in un anno le infrastrutture di ChatGPT, per un ciclo di training possono avere una impronta ecologica dalle 4 alle 200 tonnellate di emissioni di Co2 a seconda se i datacenter sono collocati in aree geografiche a basso o alto impatto energetico. Cfr. Khowaja, S.A., Khuwaja, P., Dev, K. et al.(2024) ChatGPT Needs SPADE Sustainability, PrivAcy, Digital divide, and Ethics) Evaluation: A Review. Cogn Comput (2024). Doi: 10.1007/s12559-024-10285-1.

<sup>27</sup> Si tratta della cosiddetta "promptology" o la tecnica di porre in modo efficace domande a questi strumenti. Anche in Italia vengono pubblicati manuali e prontuari di questo tenore che possono avere una loro utilità ma, a parere di chi scrive, non possono essere disgiunti da una consapevolezza più approfondita del mezzo e dei suoi limiti.

Sostenere la libertà di espressione, la privacy e la proprietà intellettuale, la sicurezza e il benessere; promuovere contenuti attenti all'uguaglianza di genere.

Le raccomandazioni<sup>28</sup> proposte nel documento Unesco per fondare il rapporto tra MIL e strumenti di intelligenza artificiale sono dieci e possono essere sintetizzate nei punti di una possibile agenda:

a) strutturare/codificare la conoscenza, rendere "spiegabile" l'IA:

elaborare un curriculum di base per una AI literacy con una prospettiva umano-centrica, che tenga conto di aspetti culturali e di uguaglianza di genere; adottare guide e progetti educativi ad hoc, anche attraverso un portale aperto e specifico per bibliotecari, educatori e trainers.

b) consapevolezza delle differenze e monitoraggio:

tenere conto del digital divide tecnologico esistente sia in termini di genere che di età; creare un osservatorio per l'IA globale in grado di produrre evidenze su come le persone apprendono/comprendono l'IA.

c) organizzarsi come istituzioni educative:

proporre un'apposita raccomandazione Unesco sull'empowerment degli utenti resa possibile dalla MIL, per una governance comprensibile e affidabile dell'IA; organizzare un evento intersettoriale presso l'UNESCO per esaminare i problemi di affidabilità delle fonti con i principali portatori di interessi di tutti i settori e ONG (tra cui l'IFLA) per garantire l'integrità delle società della conoscenza in futuro.

- *Conclusioni*

Questo documento Unesco, che adotta una visione molto "di scenario" e ampia (propria, del resto, ad un'istituzione internazionale globale), mette in atto alcune operazioni culturali che a mio avviso vanno sviluppate e fatte "atterrare" nel nostro contesto.

La prima operazione culturale è quella di considerare l'IA generativa come un nuovo media, i cui contenuti devono essere oggetto dell'analisi critica propria dell'Information literacy dal punto di vista dei criteri di verificabilità e attendibilità.

Correlato a questo aspetto mi pare importante attribuire all'Information literacy il ruolo di fornire empowerment alle persone nel relazionarsi in modo critico e informato ai contenuti prodotti dall'IA.

Un altro punto di cui tenere conto mi pare un implicito invito a sfuggire due tendenze opposte e sterili: da un lato il tecno-entusiasmo acritico verso i tool dell'IA e di conseguenza la promozione di approcci formativi di mero *adattamento passivo* a questa tecnologia, come ad esempio attività addestrative finalizzate a compulsare prompt "efficaci"; dall'altro atteggiamenti di rifiuto

---

<sup>28</sup> Cfr. UNESCO (2024), cit. p.13.

aprioristici, esclusivamente basati sui limiti dello strumento, i bias e gli errori che esso può generare.

Nel biblio-contesto che abitiamo l'Information literacy ci offre indubbiamente un robusto setting per un approccio equilibrato che preveda sia la conoscenza degli strumenti di IA e della loro rapida evoluzione che l'utilizzo cauto e vigile, ma animato da quella curiosità intellettuale senza la quale non è possibile apprendere nuove cose.

Occorre a mio parere, infine, la consapevolezza che l'IA non è una moda passeggera o "*trendy*", ma una tecnologia imponente e complessa, portatrice di interessi economici e geopolitici globali; ed è qui per restare e cambiare il nostro modo di percepire la realtà ed essere nel mondo.